28 June 2019

Submission of comments on 'Discussion paper: Use of patient disease registries for regulatory purposes – methodological and operational considerations (EMA/763513/2018)

Comments from:

| Name of organisation or individual |
| --- |
| **EFPIA, European Biopharmaceutical Enterprises and Vaccines Europe** |

*Please note that these comments and the identity of the sender will be published unless a specific justified objection is received.*

*When completed, this form should be sent to the European Medicines Agency electronically, in Word format (not PDF).*

1. General comments

| Stakeholder number*(To be completed by the Agency)* | General comment (if any) | Outcome*(To be completed by the Agency)* |
| --- | --- | --- |
|  | [INTRODUCTORY COMMENTS]We would like to applaud EMA to this guidance with its clear concept of registry and registry studies. It is a very well-written document describing considerations when planning on building new patient registries to support regulatory evaluations moving forth. EFPIA, EBE and Vaccines Europe have a vision for patient registries (health service or disease based): that they be maintained as a core part of the health information infrastructure, supporting healthcare systems to deliver quality care to patients and providing a high quality research platform for the life science sector to utilize in the discovery and development of new treatments and the optimization of existing treatments. The discussion paper is very detailed and thoughtfully outlines the concepts and methods and provides different sets of considerations (good registry practice) for patient disease registries and for registry studies. Indeed, it needs to be seen as part of a larger programme of guidance addressing RWE as a whole, together with the **Big Data Taskforce Report** and the EMA ambition for RWE outlined in the **draft Regulatory Science Strategy to 2025**.The substantial comments we provide here reflect that the Draft Reflection Paper actually opens the conversation on a number of critical themes, each of which could be the subject of a specific guidance document. Indeed, one key conclusion could be that the ultimate need is for multiple different guidance documents that are bridged by an overview paper, which could indeed be this Draft Reflection Paper. With that in view, many of the comments provided here may inform the creation and development of those supporting guidance papers. Further refining the scope of the discussion paper itself would at least help to manage expectations of the reader. The goal in the ‘Foreword’ sets out the goal to facilitate use of patient registries to support regulatory decision-making; but the paper delivers guidance on methodological and operational aspects of patient disease registries more generally and also the needs of other stakeholders. Setting a scope for the paper that matches the content, and mapping out what else needs to be considered outside of this paper, would be helpful.We would welcome the opportunity to discuss these issues and the comments herein more fully with the Task Force, and we remain available for a meeting, if that should be the determined way forward. We look forward to learning the next steps for the work and we hope that will be soon. |  |
|  | In the remainder of this document, more detailed comments and suggestions are included, as we look to address how registries are set up, governed, linked with other data sources, analysed and used in decision-making. **Our key messages** can be summarised as follows:* **Taking a more strategic view of registries and registry studies**. The paper and its readers would benefit from a more developed and comprehensive articulation of registries and registry studies, the organisations behind them, how they are established and sustained in the long run and how international alignment on standards and analysis could be achieved. This includes the broader range of decision-makers who will use registries and registry studies (including HTA bodies), as well as the underpinning data sources.
* **Governance and the life cycle of registries.** More detailed discussions are needed to clarify the requirements and means to support governance of a registry: specifically the data governance requirements now under GDPR, but also the organisational governance, including the life cycle for registries.
* **Clarifying the terms and governance related to registries and registry studies**. The draft Discussion Paper makes considerable strides in defining and setting expectations for the development and use of registries and registry studies. Understandably as this practice is still maturing, some topics still need more clarification and development of the guidance.
* **Future proofing the work**.In parallel to the development of registries, medical treatments are likewise evolving, with additional complexities that need to be considered in this guidance or at least anticipated for the future. This includes the greater impact of digital technologies and combination products.
 |  |
|  | [ADDITION TO PROPOSE]**Vision / motivation:** EMA’s overall vision of stronger collaboration between registries, sustainability of disease registries and EMA’s willingness to support interactions is only mentioned in the last paragraph of this guidance. To enable a rapid implementation of the guidance it would help all stakeholders if vision and motivation are clearly described at the very beginning of the guidance. In like manner, it will be helpful to identify the intended target audience early in the document (e.g. Introduction). |  |
|  | [\*\*ADDITION TO PROPOSE]It would be helpful to **add a section on existing registries** (including country-level registries established by medical societies) **and their potential role to support regulatory decision-making** (for EMA and other bodies)**,** so it doesn’t give the impression that they are completely out of consideration. And if they are out of consideration, refer us to other workstreams, ongoing work so we understand better the work being done at this level. **Maybe a section at the beginning recapping current status, ongoing efforts, challenges and opportunities, etc. which led to this guidance would be helpful.** |  |
|  | [ADDITION TO PROPOSE]The paper suggests the value of using registries as the basis for registry studies (data is already collected, ability to compare across treatments, standardized data, time). It may be helpful to draft a section setting out explicitly the strengths/advantages of registries (including specific types of registry), including the limitations of setting up registries and indeed limitations in their use. Some types of registry have methodological and operational specificities that are not mentioned in the document: Pregnancy registry, Paediatric registry, Rare (or ultra-rare) disease registry and Cancer registry. We acknowledge that guidelines already exist detailing these specificities. The guidance should direct the readers to these references to manage scope and expectations on these specific cases. |  |
|  | [\*\*\*ADDITION TO PROPOSE]Some language on the benefit to the various stakeholders (MAH, regulators, physicians, patients, HTA bodies, registry coordinators…) may be useful and would support the sustainability of registriesAt the EMA- EuropaBio Info Day (22.11.16) Dr. Alison Cave (EMA disclaimer included) pointed out that 64% of the registries analysed by Bouvy et al. did not collect any HTA related variables (slide 13). This was specifically criticised under duplication of efforts and lack of sustainability. This could be better developed in the Discussion Paper, including for example the core data elements, which currently solely focus on the perspective of regulatory registry studies, even though the chapter is on good registry practice (so overarching all registries aiming to increase sustainability). We believe that it would be **helpful to either extend this list with HTA/QoL related core data elements or list it as regulatory relevant core data elements**. Otherwise, by using this list the EMA is contradicting themselves in the main message of registries not being sustainable and needing to be re-used/used for multiple purposes. |  |
|  | [ADDITION TO PROPOSE]Please consider including some discussion on assessing patient diversity. |  |
|  | [\*\*\*ADDITION TO PROPOSE]Considerations regarding the use of registries in the era of digitalisation of medical health records could be added and cross-referenced to the HMA/EMA Big Data Roadmap, when this is further defined.At several points in this document**, the linkage of existing databases/ registries is encouraged** (e.g. page 16, line 20-22). It would be very helpful to understand how to best link these databases into **“central/umbrella registries”** (including e.g. selection of databases, unique identifiers, data dictionary creations, etc.). This should include recommendations on the access of registries by independent researchers, with the potential to link these data to others not pertaining to the registry itself. One area of increasing interest is the pooling/combining of data from multiple distinct registries or the linkage of registry data with other data sources (e.g., administrative claims, EHRs). This can potentially allow for a richer and more complete source of patient data. This is very briefly touched upon in this discussion paper when discussing other issues (e.g., page 34, lines 14-15). However, more specific methodological and operational considerations for data pooling/linkage, either when the data sources already exist or when a registry is being actively designed, would be beneficial for the audience. Ideally, there would be an objective to create European registries which fulfill requirements in EMA and HTA processes.Clearer guidance/push could be given to the “one registry approach” and how different stakeholders should try and work on setting up one multipurpose registry instead of several ones (especially in rare disease). The document could also clarify further new data types such as “omics”, digital health sources challenges and opportunities. Discussion of pooling different registries and which disease benefit the most could be includedEMA could also offer guidance on how to facilitate discussion on existing registries in the same disease area – which may entail several orphan diseases. |  |
|  | [\*\*ADDITION TO PROPOSE]**Timing and Timelines:** The ease of implementing suggestions in the document depends on early interactions between the registry system group and the potential registry study stakeholders. It might be helpful to clarify the importance of this up front.In the Patient registry initiative-strategy and mandate of the cross committee task force (EMA/180341/2017) objectives 4.1.a specifies registry timeline objectives were set as *“****embed registry questions pre-submission/ validation, day 120, and day 180 time points*”.** Why have they not been included in this guidance and can the timelines be again specified? |  |
|  | [\*\*ADDITION TO PROPOSE]**Joint registry study with multiple MAHs.** Multiple MAHs participating in a joint registry study based on a single protocol as encouraged in this guidance would, in case of a PASS, each hold sponsor obligations (page 8, line 5-6; page 36, line 34-37). It would be appreciated to receive guidance on the how to split obligations, particularly on the governance and sponsor responsibilities. One alternative would be for the regulatory body to be the ‘sponsor’ of a comparative PASS, propose a protocol and SAP and ask for comments from all relevant MAHs for transparency. |  |
|  | [ADDITION TO PROPOSE]The concept of “**Good Registry Practice**” could be even more substantiated in this guidance, at least to set out an international ambition for this work. Obviously the EU institutions are only addressing goals within the EU, but the ambitions must be for **international standards** **to be developed**. This could be **an ICH goal** and indeed would be valuable proposal for a topic, with this guidance once finalized as an important first contribution.  |  |
|  | [\*\*ADDITION TO PROPOSE]Recognising that this is technical guidance, the document says very little (only p43 lines 35 – 40) on one of the most defining aspects of registries: **funding**. Although that may be beyond the recommendations scope, it should be more fully addressed as a determining factor in addressing the recommendations for good registry practice. E.g. exhaustive enrolment is somewhat determined by the resources available to the registry team. Section 5.8 on Governance really doesn’t address this issue directly, including its implications.It also leads to another aspect of registries not addressed – **what should happen when they are closed**. What is the right way to close a registry? What would be the possibilities for archiving a registry to preserve the investment of patients and others for research, recognizing that this could potentially be judged a breach of a GDPR obligation, i.e. to retain personalized information only for as long as it is required for a fair and legitimate purpose’. |  |
|  | [\*\*ADDITION TO PROPOSE]Although the title of the Discussion Report is “Use of patient disease registries for regulatory purposes – methodological and operational considerations”, it is in line with both stated EMA current policy and the **draft Regulatory Science Strategy** that evidence development for regulatory purposes must also be considered in light of decision making along the path to the patient – **namely HTA bodies**, that are also interested to use results in their assessments. This should be more specifically addressed and clarified in the Discussion Report. |  |
|  | [\*\*\*ADDITION TO PROPOSE]Propose considering potential for alignment with International Medical Device Regulators Forum (IMDRF) output, e.g. policies like [Integrating patient registries and innovative tools for enhanced medical device evaluation and tracking](http://www.imdrf.org/workitems/wi-patient-registries.asp), including its [recent consultation](http://www.imdrf.org/docs/imdrf/final/consultations/imdrf-cons-registries-n46-pd1-170817.pdf) on these tools. This is especially relevant where there is a preference for disease rather than product-based registries, as the former could include patients treated with medical devices and all data collected with medical devices.  |  |
|  | [ADDITION TO PROPOSE]**Suggest adding of a glossary of terms** It will be helpful to add a separate glossary or extend section 3 with definitions of key terms used in this guidance document. Inclusion of specific examples will enhance clarity of the document. |  |
|  | [CONCEPTUAL CLARITY]**Differentiation between primary data collection and secondary use of existing registry data.*** A **clear understanding of primary data collection as opposed to the secondary use of existing data** should be made in the context of a registry study in an existing registry or newly initiated registry. In the guidance, it is unclear what the conditions to use one versus the other are. The document may benefit from examples of this definition.
* The differentiation between primary data collection and secondary use of existing registry data is highly important for this paper. However, in some sections this is not clearly differentiated though (e.g. Data quality page 8). It would be important to have this clearly indicated.
* One suggested approach to differentiation:
	+ proactive use of existing data
	+ proactive initiation of a registry / collection of data
	+ Retroactive use of existing data (e.g. answering Agency questions)
 |  |
|  | [\*\*CONCEPTUAL CLARITY]**Interventional vs. non-interventional registry studies** Through the introduction of Core data Elements, multi-national registry studies could be considered interventional in some countries and non-interventional in others, caused by differences in the routine clinical practice. The classification depends on local implementation of the EU directive (2001/20/EC). Moreover, ‘low interventional’ is only relevant under the CT Regulation, which is not yet implemented. In multinational studies, we may have different classifications: interventional, low-interventional and non-interventional. How are these regional differences supposed to be addressed?  |  |
|  | [CONCEPTUAL CLARITY]Quality management is a key theme, but there’s little mention of what is a “high quality” registry - more details to understand what quality should be, with measurable criteria, would be welcome (e.g. page 24, what is the % of completeness, or range of acceptability of completeness?).Although a lot of points to consider are provided, it may need to go further in order to clarify expectations and avoid different interpretations of what is quality and standardisation. There is no mention of geographical spread/diversity either and evidence quality. . |  |
|  | [\*\*CONCEPTUAL CLARITY]Different stakeholders have different purposes for data collection (e.g. EMA and HTA bodies); this may impact the aggregation level and information requirements which ultimately will lead to variability in the data collection process. Therefore, **stating clear aims for a registry is important.** Or providing minimum requirements to be recorded in the source data could be considered, so depending on aim of the study, data extraction can be adapted. We can envisage a certification/ qualification procedure ultimately. |  |
|  | [CONCEPTUAL CLARITY]There are several mentions of the ADVANCE Code of Conduct “for vaccines” – although initially developed indeed as part of a Vaccines IMI protect, it could (should?) apply to any medicinal product. Suggest deleting “for vaccines” when the reference to this CoC is made. |  |
|  | [CONCEPTUAL CLARITY]**In Section 5.8 Governance,** should be more explicitly linked as the main factor determining sustainable management of the registry. This could then be occasion in the guidance to address the circumstances whereby the registry holder is no longer able to maintain the registry. This is closely linked to the impact of sustainable funding for a registry, and therefore effective governance will also need to address this practical but pivotal factor. |  |
|  | [CONCEPTUAL CLARITY]Definitions of the various registries could **refer to existing reference texts** (see also comment above), and how registry studies relate to epidemiological studies (e.g. in terms of design) could be clarified (the guidance effectively speaks about methods of epidemiological research). |  |
|  | [\*\*\*CONCEPTUAL CLARITY]Please provide clarity on the **requirement for collection and reporting of ICSRs from registries and also from ‘registry studies’** that are:* Funded by academia/medical research agencies alone.
* Conducted by academia/medical research agencies with funding from a MAH (particularly where the MAH is conducting a ‘registry study’ using primary or secondary data from such a registry).
* Wholly conducted by a MAH.
* Which are focused on a disease area for which an MAH has an authorisation in the EEA, but where the product is not the focus of the registry (the focus of this Paper).

Clarity is requested regarding whether ICSRs should be considered ‘solicited’ in such circumstances because the MAH has an authorised treatment for the disease in question. Clarity is require regarding whether there are situations in which spontaneous ICSRs might be appropriate.There is a potential for duplication of reporting of ICSRs where reconciliation is not feasible. The reporting obligation is clear where an MAH conducts a registry study using primary data collection. In other contexts, the ‘organiser’ has this role and the MAH is informed by the regulatory body.It is worth recalling that **many current registries are partially funded by multiple MAHs**. It is common practice for disease registries to be managed by non-profit organisations with several MAHs funding / supporting / consulting the registry.The current proposal would not be workable for the registry owner or for the MAH due to the safety reporting requirements. |  |
|  | [CONCEPTUAL CLARITY]The safety requirements are vague and not consistent with the GVP modules. Please ensure that the safety reporting requirements throughout the document are clearly described and fully aligned with the GVP Modules 6 and 8, or alternatively refer to the GVP modules.  |  |
|  | [CONCEPTUAL CLARITY]**Umbrella registries** In several citations (e.g. the AHRQ guidance on registries) and a multitude of contexts the combination of multiple national registries into one overarching registry is called umbrella registry (e.g. TREAT-NMD, BOLD, RD-HUB and many more). This guidance uses the term “central registry” on page 19, line 14, would this be considered the same as “umbrella registries? For the clarity of this guidance, it would help to use a defined term for these “umbrella/ central registries”. This would be particularly helpful in the context of harmonizing registries (e.g. page 24, line 33-40).***Proposed change:*** *use a specific term such as “umbrella registry”* |  |
|  | [EDITORIAL, TERMINOLOGY]**Patient disease registry vs disease registry:** In the definitions, it is explained that patient registries can be subdivided into either e.g. disease registries or product registries. However, the combination of terms is unclear, at times they are referred to patient disease registries (page 1, line 8; page 2, line 5; page 3, line 10; page 6, line 2-3/ 9/ 19/ 20/ 26/ 27; page 10, line 17; page 11, line 2-3/ 7; page 12, line 36; page 16, line 1/ 15; page 23, line 30; page 30, line 19/ 21-22/ 36; page 36, line 21), other times they are only referred to as “just” disease registry (page 3, line 8; page 6, line 17; page 7, line 9/ 14; page 9, line 21; page 10, line 21/ 24; page 12, line 32; page 13, line 3/ 5; page 14, line 1; page 20, line 18; page 22, line 1; page 28, line 3/ 24/ 25/ 26; page 30, line 32; page 31, line 4/ 32; page 32, line 16/ 22/39; page 43, line 4/ 35/ 41). Product registries are never referred to as patient product registries and always “just” described as product registries. Consistent use of terminology and/ or clarification of overlap and/or differences would be helpful. As a result, it is not clear what type of registry is/ are meant when “just” registry or patient registry is used and whether both disease and product registries are spanned under this term, considering the title of the guidance refers to patient disease registries and only population registries are excluded. ***Proposed change:*** *we would propose to use consistently disease registry where applicable* |  |
|  | [EDITORIAL, TERMINOLOGY]**Registry study definition.** There is no definition of “Registry study” in section 3.1 while bits and pieces are given throughout the guidance. The guidance would benefit from one main definition in Section 6. Furthermore, the scope of registry studies should be clarified. Even if only registry studies mandated by Regulators were considered, in the paper by Bouvy et al. (2017; DOI: 10.1002/pds.4196), the primary objectives of these mandated registries were either safety, or pregnancy safety, or efficacy/effectiveness, or disease epidemiology. Therefore, the entire regulatory context/ all mandated registry studies cannot be described by “just” PASS and PAES (as seen e.g. page 7, line 28-29). Drug utilization studies are mentioned page 16 but no information is provided in Part 6 where only PASS and PAES are detailed.The definition could also consider database linkage studies using publicly owned databases, such as the Nordic national databases.***Proposed change:*** *include a definition of registry studies e.g.**“Detailed investigation and analysis of a research question or hypothesis using a registry as a source population” (as presented by Xavier Kurz at the Big MS Data Group Consortium 22 February 2019)*  |  |
|  | [EDITORIAL, TERMINOLOGY]In Section6, the responsibilities of the MAH and registry coordinator are not clearly defined. It could be helpful to have a table summarizing these.  |  |
|  | [EDITORIAL, TERMINOLOGY]This paper uses both **secondary data collection** (page 7, line 43; page 8, line 16; page 14, line 13; table 1; page 16, line 19; page 36, line 30; page 41, line 28) and **secondary use of data** (page 9, line 6; page 14, line 20; page 40, line 43). For better understanding, we suggest replacing the term “secondary data collection” by “secondary use of data”, as registry data when used for a registry-based study are not really “collected” for the study, but more extracted from the registry and then used for the study. In addition, this would be better aligned with the way it is also phrased in the GVP modules (especially VI and VIII).Consistent use of terminology and/ or clarification of overlap and/or differences would be helpful. The GVP (e.g. EMA/813938/2011 Rev 3\*) only refers to it as secondary use of data. We would propose changes to be aligned with the GVP terminology. |  |
|  | [EDITORIAL, TERMINOLOGY]**Primary data collection vs. active data collection**Would active data collection only refer to primary data collections meaning the collection of prospective data or include “secondary data collection” even though, only retrospective data is utilized. Please clarify the overlaps and/or differences between active data collection (page 7, line 12; page 14, line 3; table 1; page 28, line 9) and primary data collection (page 7, line 44; page 14, line 13; table 1; page 16, line 17-18; page 36, line 31; page 38, line 19; page 40, line 41; page 41, line 3/ 35).***Proposed change:*** *If the intent is similar and both could be understood interchangeably, we would propose to use the same wording throughout the document, primary data collection.* **Primary data collection for registry vs. natural history** Discussion of the limitations of registry data in comparison to collection of natural history and untreated patients should be developed further. In particular, gguidance on the timing of setting up a registry would be helpful, outlining the value of the collection of natural history as much as possible before available treatments / procedures are available.**Primary data collection definition**This paper **defines primary data collection** inconsistently as either *“where the events of interest for the study are collected directly from patients as they come to the attention of the investigator”* (page 7, line 44 - page 8, line 1; page 16, line 17-19; page 36, line 31-32) or *“…collection directly from HCPs or consumers (i.e. where the events of interest for the study are collected directly from patients as they come to the attention of the investigator), …”* (page 9, line 4-5; page 40, line 41). Consistent use and definition of terminology would be helpful.**Proposed change: “…***where the events of interest for the study are collected directly from the patient, caregiver or other healthcare professional as they come to the attention of the investigator.”* |  |
|  | [EDITORIAL, TERMINOLOGY]**Missing data vs. non-available data**The differences between missing data and non-available data should be clarified in the document. For disease registries the data generated globally will collect the real-world evidence of which assessments patients are undergoing in different parts of the world. Which does mean that for some assessments the data will not be available for analysis.Definition suggestion:*Missing data:* Data that has been generated by the treating physician and hence is available for data entry in the data collection system – can be managed and “controlled” through site management*Non-available data:* data that has not been generated (as does not follow the standard clinical care at the treating physician due to different treatment guidelines and or country regulations and or insurance considerations) – data will not be available for data entry*Core data***:** ensure that this is not perceived as a “core data set that every patient has to complete (as this would interfere with the observational |  |
|  | [EDITORIAL, TERMINOLOGY]Using the term “investigators” in registry-based non-interventional studies may be misleading. The regulation talks about HCP which would seem to be more appropriate in this setting. |  |
|  | [EDITORIAL, TERMINOLOGY]We acknowledge that the landscape of recommendation and existing documentation on registries and registry studies is complex. The guidance will benefit from **precise reference to external documents rather than rephrasing** to avoid confusion. Some example are highlighted below but might not be exhaustive:* GVP Module VI and VIII : This document could explain or at least reference specific GVP requirements for collection, analysis, and reporting of adverse events in registries and registry studies. It is stated in the current draft of the document (e.g., page 14, lines 12-14) that, “For registry studies, requirements to MAH for suspected adverse reactions differ between studies with a design based on primary or secondary data collection”. However, the nature of this difference is not explained. At the very least, references to relevant sections of GVP modules VI and VIII should be provided. We nevertheless recognise that not all registry studies are PASS.
* GVP module VI C. 1.2.1.2, submission of ISCRs is not required in studies with secondary use of data, but “all adverse events / reactions collected for the study should be recorded and summarized in the interim safety analysis and in the final study report unless the protocol provides for a different reporting with due justification”. This wording could be interpreted as implying that aggregate reporting and analysis of all adverse events is required in registry studies with secondary use of data. However, according to GVP module VIII, B.3.1. “For studies based on secondary use of data, a statement should indicate if adverse events / adverse reactions are analysed”, which implies that such analysis is not mandatory.
* Additional insights on the qualification process by the CHMP would benefit the document and the different parties. In light of already performed qualifications (ECFPSR and EBMT), the list of minimal requirements for a registry to be eligible for health authorities/regulatory assessments/evaluations would improve clarity within the guidance and also facilitate interactions between MAH and registry coordinators. It would furthermore help registries in their willingness to improve and achieve sufficient quality standards if needed.
* Gold standard public inventory of registries: In this guidance public inventory of registries are referred to at several points. In previous documentations of the Patient Registry Initiative (e.g. (EMA/69716/2017; EMA/180341/2017), the PARENT-JA Registry of Registries was given as the gold standard/ tool to use. Is this still applicable, which public inventory of registries should be used best?
 |  |
|  | [EDITORIAL, TERMINOLOGY]The guidance uses “safety reporting” as term for safety case reporting and safety monitoring of an adverse event of special interest. That raises confusion. The guidance would gain from a distinction **between “safety case reporting” under registry and “safety monitoring” under registry study**. |  |
|  | [POTENTIAL CONFLICT IN PURPOSE]For disease registries, safety reporting according to national requirements is proposed. For registry studies, GVP will apply. With respect to GVP requirements around all AE reporting, this may be logistically challenging for registries, and ability to provide this for registry studies, when registries were not initially designed with this in mind, may be difficult.As noted in this response, not all registry studies meet GVP VIII criteria as PASS, ie if the main objective is not safety related. |  |
|  | [POTENTIAL CONFLICT IN PURPOSE]There are several mentions of PAES. This may be restrictive, as PAES follow a strictly defined regulatory framework. For some medicinal products (such as vaccines), additional (post-licensure) efficacy studies are not expected to provide substantial new information or are not feasible, whereas observational effectiveness/impact studies are required. Suggest using the broader term of effectiveness, or PAES/effectiveness. |  |
|  | [POTENTIAL CONFLICT IN PURPOSE]**Analyses of data in MAH sponsored PDC registry.** Who is supposed to do the analyses? How is Pharma supposed to not have access to the data but still fulfil sponsor obligations such as oversight over data? It would be helpful if it is defined what should be covered in contracts with independent third parties in a registry study. |  |
|  | [SCOPE]The remit/focus of the document – patient disease registries – is clearly outlined. However, **several other types of registries are used for regulatory purposes** (exposure, drug), which have similar challenges, and can sometimes be used in combination if linkable. A rationale for the focus on disease registries is provided on page 13 (lines 1-9), however, the importance of other registries could be clarified and/or reference to other guidance documents or key publications describing registries in general could be provided. |  |
|  | [SCOPE]In multiple places, the importance of safety reporting in Registries and Registry-based studies is emphasized, including those registries that are wholly independent of any MAH. While we agree in principle that safety reporting is paramount, in practice we have seen that the requirement to report ALL non-serious as well as serious AEs can be a significant barrier to participation and increasingly non-naturalistic. We would advocate that the **safety reporting obligations should be “fit for purpose” and ideally limited to the potential AEs of greatest concern** to maximize recruitment and retention. In particular, most non-serious AEs are unknown and unreported to the registry clinician and are thus of uncertain value in determining the safety profile of a therapy. Collection, reporting, verification, and reconciliation of AEs that provide little to no new information on the safety profile of a product divert resources from other important activities and information needs. |  |
|  | **Stakeholders:** * **HTA’s**In the Patient registry initiative-strategy and mandate of the cross-committee task force (EMA/180341/2017) objective (4.1.b) include a list of stakeholders working jointly towards core data elements including e.g. Reimbursement bodies and to overall increase sustainability of registries. Many of these stakeholders are not consistently considered in Section 5 Good Registry Practice and the focus solely remains on registry use for regulatory purposes in this guidance.
* **National experts/ KOLs / Patient groups**As part of the Patient Registries Workshop, 28 October 2016, Observations and recommendations arising from the workshop (EMA/69716/2017) in chapter 2.1. Chronology of EMA Registry Initiative activities it is recommended that the input from national experts/ KOLs is sought out. It would be helpful to learn more about the format, content and platform for seeking this input could be established as part of this guidance. Focused efforts should be supported to involve patients (associations), e.g. via learning from existing successful examples.
* We recommend to host a cross-functional discussion of this paper should be ensured to also have input from various stakeholders, e.g. physicians, HTA bodies, to improve future conduct of registries and broad acceptance of their results and – no less importantly - to avoid duplication of efforts (due to different requirements). Experience in the past has shown that motivation of physicians/ hospitals to participate is also key to secure timely and effective data collection.
 |  |

1. Specific comments on text

| Line number(s) of the relevant text*(e.g. Lines 20-23)* | Stakeholder number*(To be completed by the Agency)* | Comment and rationale; proposed changes*(If changes to the wording are suggested, they should be highlighted using 'track changes')* | Outcome*(To be completed by the Agency)* |
| --- | --- | --- | --- |
| **Abbreviations**Pg 5 |  | The acronym SNOMED should be defined. Also, it is proposed that definition of the following roles which are references within the documents (page 25) be included:* Registry coordinator
* Local Registry coordinator
* Data custodians
* Data manager

Proposed change (if any): Add definitions of the above roles here or within the documents. |  |
| **EXECUTIVE SUMMARY** |
| **Executive summary**Pg 6, lines 10-11 |  | Patients registries could also be established by pharmaceutical companies in certain circumstances (e.g. to address a post-licensure commitment for a new drug, for rare disease, etc), with the support of academics, patients and other stakeholders. This possibility is also mentioned on Page 7 (lines 12-14). The drafting also **misses that oftentimes these registries are established as a collaboration**, across organisational types and geographies.*Proposed change:* ***“****Patient disease registries may be established by public organisations such as academia or medical research association of health care professionals or patients, by industry or by collaborations across types of organisations and across geographies.”* |  |
| **Executive summary**P6 line 12 |  | Should clarify that registries are also used to monitor effectiveness, more readily than efficacy. Effectiveness is more commonly used in real world setting. *Proposed change:* “..to monitor the (~~efficacy~~), safety and/or effectiveness of treatments”  |  |
| **Executive summary**Pg 6 line 17 |  | “regulators generally prefer disease registries (…) to product registries”. The “generally” does not make it a universal rule, but it could be helpful to acknowledge some situations in which product registries can be preferred, in some circumstances initiating a treatment (not necessarily by a specific drug, but with a class of drugs) can be a marker of disease severity that could be useful to restrict the population to include in the registry as alluded in section 5.2 lines 35-36, in some situations (e.g. diagnosis codes not specific enough for some rare diseases) products received may be a more reliable way to identify the eligible population |  |
| **Executive summary**Pg 6 lines 19 - 40 |  | It may be appropriate to introduce the possibility of multiple registries earlier in the text e.g. following Page 6, line 19 (see below). Furthermore, in the USA, MDEpiNet (Medical Device Epidemiology Network) uses the term “coordinated registry networks,” to capture the concept of ensuring some basic compatibility across registries. Consideration to adopting this terminology could also be given. We note that references to examples of where core data elements are provided appear on Page 20, lines 18-24. Perhaps these can be mentioned at a high level in the Executive Summary. Proposed change: *Line 19:* controlled designs without an external data source. I**n some instances multiple patient/disease registries may exist (note the difference from a single registry at multiple sites).** The main focus … |  |
| **Executive summary**Pg 6, lines 20-21 |  | The main focus of this document is clear except when it is mentioned that the considerations of the document also apply to product registries or registries of patients defined by a specific condition. In this lateststatement, what is covered by “defined by a specific condition” may be not so clear, as a specific disease (e.g., Multiple Sclerosis) could be considered as a “specific condition” and is then not different from the conditions of a patient disease registry. It might also vary not only by pregnancy or age but also a specific disease or a stage of a disease could be considered as a “specific condition”. *Proposed change*: It would be appreciated if the Agency could give some examples of specific conditions which would trigger the set-up of a registry without being considered as patient disease registry. Or to delete the second part of the sentence after “product registry” which would limit the possible confusion. |  |
| **Executive summary**Pg 6, lines 25-26 |  | The reference to work in **accordance with GDPR** does not address the real concerns that emerge in the details of using registry data. **A clearer mapping of the potential challenges should be included** (particularly in section 5.8.3) |  |
| **Executive summary**Pg 6, lines 29-30 |  | This refers to “i.e. the situation where enrolment is influenced by patient characteristics that may affect the validity of the analyses”. However, the analyses may be “valid” (in that they accurately address the question as it is defined), but they may be addressing the wrong population. It should be clarified whether the text is intended to address the possibility that the outcome assessment may not be representative of the outcomes in the target population of interest (if, for example, only patients with more favorable prognoses are included in the registry.)  |  |
| **Executive summary**Pg 6, lines 28-35 |  | Please clarify the scope of the registries/registry studies that are required to be registered. Is it the ones in scope for PASS /PAES, ie. required by EU regulators? |  |
| **Executive summary**Pg 6, line 32 |  | It is stated that the MAH should ensure the study is registered in a public database. However, it is not clear which database is suited for such registrations of retrospective data collection/secondary use of date for ‘Registers’ and/or ‘Register studies’. |  |
| **Executive summary**Pg 6, line 34 |  | The **time elements** seem to be limited to the “important events**”. This time information should not be limited to the “events” but also to the different assessments done**, **as this may be critical.** As an example, in oncology, not only the progression date should be recorded, but also the dates of all the previous examinations performed up to the progression to know if the duration of response is linked to one assessment only (during which progression was observed) or regular assessments up to the one during which the progression was observed. Same for the treatments which should be recorded with clear start date and end date.*Proposed change*: to expand the time elements beyond just the important events. |  |
| **Executive summary**Pg 6, lines 37-40 |  | ***Core data elements***: a list of core data elements to be collected in all patients is proposed. They should be **harmonised or mapped across registries for a same disease** to support regulatory evaluations and facilitate implementation of a common data quality system, data exchange, common data analysis and interpretation of results from different registries **i.e. coordinated registry networks**.  |  |
| **Executive summary**Pg 7, lines 3-8 |  | The quality management is a key topic for registries, as for any other data sources. And this is not a “one-time” effort, but more an iterative process which has to be performed regularly. It could be worth to emphasis this point as early as in the executive summary.*Proposed change*: To highlight in this section already that the quality management should be a constant and iterative effort of the registry coordinators. |  |
| **Executive summary**Pg 7, line 11 |  | *“Reporting of suspected adverse reactions through the national or regional pharmacovigilance system should be encouraged”.* Is encouraged strong enough wording? **Shouldn’t it be required**? This occurs in some other instances in the document e.g. **Pg 8, line 22**, recommending documenting difference between incident and prevalent cases rather than requiring this, and **Pg 8 line 37** where auditing *may* be needed rather than *required*. We would anticipate that this would be only where required by regulatory authorities (eg if the registry meets the definition of PASS). |  |
| Pg 7, line 9-18 |  | *Safety analysis*: disease registries conducted by organisations such as academia or medical research associations should follow the national requirements for the management of safety data. Reporting of suspected adverse reactions through the national or regional pharmacovigilance system should be encouraged. “Any active data collection system put in place in a disease registry and initiated, managed or funded by a MAH to collect and record suspected adverse reactions to one of its medicinal products must follow the regulatory framework for PASS.” **Disease registries are generally not suitable for a rapid statistical analysis** **of new safety signals** but they may be useful for the monitoring and characterisation of known or suspected adverse reactions. A list of adverse events of special interest (AESI) can be defined and integrated in the routine data collection system. **Proposed change:** Please include additional information on how to deal with/interpret the influence of co-medications due to co-morbidities.Please consider noting that registries may not be suitable for rapid statistical analysis of new safety signals, and why this is the case.  |  |
| **Executive summary**Pg 7, lines 12-14 |  | This needs to be clarified, as not all registries set up by a MAH will meet the GVP Module VIII criteria for being a PASS. The collection of safety data alone does not mean that a registry is a PASS, unless the ‘main aim’ of the registry is related to safety (or one of the other criteria established in GVP Module VIII.)*Proposed change:* **Please clarify the statement to ensure that PASS are correctly identified**. |  |
| **Executive summary**Pg 7, lines 17-18 |  | Include more detail as to what is intended by “integrated into the routine data collection”. For example, specific case-report forms for specific events, meant to capture an appropriate level of clinical detail. |  |
| **Executive summary**Pg 7, lines 25-26 |  | The term “data ownership” may be problematic; at issue here is mainly the “data processing” and how we ensure that any data processing within the context of registries meets the GDPR requirements. Reference should also be made to data access. The registry “owner” may “own” the data but may still permit access in some form by outside parties, presumably under a signed data use agreement. Furthermore, clarification regarding the meaning of “data sharing” is requested.Also, consent may not be the only legal basis; further consultation is required. *Proposed change*: Principles of data processing, **data access,** informed consent (where applicable) and data security should be applied in accordance with the General Data Protection Regulation (GDPR). |  |
| **Executive summary**Pg 7, lines 29-30 |  | MAHs hold responsibility for supervision of PASS and PAES: this goes beyond supervision, monitoring data etc. The MAH should be significantly involved in study design, as the owner of the commitment and according to its obligation to report results to regulators. |  |
| **Executive summary**Pg 7, lines 39-41 |  | One additional key element for the study for the sites to assess if they can participate in a study or not is not only linked to data availability and data quality requirements, but also study timelines. Indeed, some data can not be available yet when the study is discussed (especially when linked to a new product to be marketed) but could be made available within reasonable timelines which will fit the study timelines (especially when there is a need for additional data collection).*Proposed change*: to include also as a possible condition to participate or not in a study the study timelines. |  |
| **Executive summary**Pg 7, lines 42-44 |  | The need for additional data collection may not be limited to the events of interest but can also include some key confounders or variables which may be specific to the events of interests, and then not recorded routinely in the registry. Of course, if these confounders or variables need to be added frequently, then it may be appropriate to consider amending the critical data points.Please provide clarity on requirements for studies that involve both primary and secondary data collection within the same protocol.The insertion of examples of primary and secondary data collection may be helpfulProposed change: to add “where the events of interest or variables specific for the study…”And to add (**in bold**)an early decision to be made is the choice of the data collection method: secondary data collection, where the data for the study are already available and extracted from a dataset **(i.e. a registry),** or primary data collection **(i.e. a prospective observational study, using a registry to recruit)** |  |
| **Executive summary**Pg 8, lines 1-2 |  | The implications for safety reporting linked to the fact that some studies may need additional data collection, specifically for the purpose for the registry-based study, should be summarized in the executive summary as this is a critical point to be understood by all parties to be involved in a study following this rule.*Proposed change*: to include a summary of the implications for safety reporting within the executive summary. |  |
| **Executive summary**Pg 8, lines 3-9 |  | A **master protocol** is critical to ensure that the overarching objective of a multi-site registry is met.Besides GVP Module VIII which applies to PASS, other accepted good practice guidelines could be mentioned (GEP, GPP). |  |
| **Executive summary**Pg 8, lines 5-6 |  | **The sentence “For studies addressing several products, all concerned MAHs should participate in a joint registry study based on a single protocol” is misleading** as a study can be comparative, then assessing the safety or effectiveness of one product compared to another or several others. This does not mean that it is a joint registry study. This would be the case only if the concern to be assessed is for several products or a class of products and will fit with the definition of a joint PASS or PAES.*Proposed change*: Please rephrase to clarify the context of joint study.Obligation to participate in a joint registry study impedes contract negotiations between MAHs. Should be encouraged, not obliged (in line with Line 26 page 30 and Line 37 page 36).*Proposed change*: For studies addressing several products, all concerned MAHs should **be encouraged to** participate in a joint registry study based on a single protocol. |  |
| **Executive summary**Pg 8, lines 6-7 |  | The sentence “If a registry study is to be conducted across multiple sites…” can also introduce some confusion. Indeed, many registries will collect data on more than one site. And this will be done on the same way across multiple sites. So here, based on the end of the sentence (“of each registry”), it would be more correct to refer to studies to be conducted across multiple registries, within a same disease.*Proposed change*: Replace “across sites” by “across registries dealing with the same disease”, or wording consistent with “*In some instances multiple patient/disease registries may exist (note the difference from a single registry at multiple sites).* “ {as suggested earlier} |  |
| **Executive summary**Pg 8, line 11 |  | Instead of the use of conservative assumptions, the sample size estimation should use realistic assumptions, in order to ensure realistic timelines. Is the intent here to outline the need for adequately powered registries? If so, this could be more clearly stated.*Proposed change*: Replace “conservative” with “realistic”.IF the decision is to use “conservative”, then please define as this has different interpretations according to the reader. |  |
| **Executive summary**Pg 8, line 12 |  | Not only the number of patients and the duration of follow-up are key elements to consider, **but also drop-out rate (from the registry) will impact the final sample size** to be included in the study to achieve the study objectives.*Proposed change*: Please consider to add also in the list of key assumptions to consider the drop-out rates when relevant. |  |
| **Executive summary**Pg 8, lines 14-26 |  | The discussion of incident and prevalent patients seems too limited and may need to be refocused or expanded. Often, new drugs entering the market are NOT used as first line therapies, so they will not be used by incident patients. In these situations, incident user designs (also called “new user” designs) may be most appropriate, with appropriate attention in the analysis paid to history prior to starting the new drug. It is important **to distinguish incident users from prevalent users** when many registries traditionally only measure covariates at enrolment, so that prevalent users’ baseline disease severity etc already reflect treatment effect. **Potential selection bias is possible** for some new users who experience acute adverse events and not enrol to prospective registries. Methodologically, the issue of mis-measured “baseline” variables for prevalent users may be resolved by retrospectively collect those variables when they initiated the exposure (i.e., at the time when they were incident users). NOTE: ON PAGE 12, there is a more general definition of incident and prevalent, which includes incident use vs. prevalent use of a medication. Consideration could be given to aligning these two sections.  |  |
| **Executive summary**Pg 8, lines 14-16 |  | Comment: An already existing registry is a secondary data source.*Proposed change: As for other secondary data sources, methodological challenges to define study population apply also to existing registries.* |  |
| **Executive summary**Pg 8, lines 18-20 |  | The definition of incident and prevalent patients is misleading as mixing both the disease aspect and the treatment aspect. **The disease diagnosis and the initiation of a new treatment can be 2 different events**, and patients can be incident patients (newly diagnosed) or prevalent patients (not newly diagnosed), and then also incident or prevalent based on the initiation of a new treatment or not. So it would be worth to precise that then the effects (beneficial or adverse) of a new treatment need to be monitored (as written on line 16), then the focus should be on the patients for which the treatment has been newly initiated (then defining the incident patients), and less on the newly diagnosed patients, which could then be included as a covariate. In addition, prevalent patients should not be defined as “patients already included in the registry”, which would mean already diagnosed, but more as patients for which the treatment was initiated before the inclusion in the study.*Proposed change*: Clarify in this context the definition of incident and prevalent patients, to limit this definition to the new initiation or not of the treatment(s) to be assessed. |  |
| **Executive summary**Pg 8, lines 25-26 |  | Regarding the sentence on the patients having been involved in clinical trials to be enrolled in a disease registry later on, it would be worth to clarify what is meant here by “later on”. Does it mean e.g., at the end of the trial? Which could be an issue in some diseases (such as rare diseases, cancers with high mortality rates). In addition, these patients could be of value to answer some research questions, possibly not linked to drugs, and not including these patients when they are diagnosed would bias potentially the registry by excluding some sub-groups of interest as written in the document (e.g., with specific genetic variants). Finally, some patients included in registries could also later be involved in clinical trials. And then, how this should be handled? We think that above all, before excluding from the registry some patients because of their participation in a clinical trial, this is more valuable to ensure that they are included in the registry and that there is a flag which allow to identify these patients during a registry-based study, so that, based on the research question, we can check if the fact to have included these patients in the study have influenced the results in one way or another.*Proposed change*: to adapt the sentence not to say that these patients could also be enrolled in a disease registry later on but **ensure that these patients are included in the registry and easily identifiable** so that the impact of their inclusion in a study when this is the case can be assessed. Suggested edit: “… could also be enrolled in a disease registry **post-clinical trial** ~~later on~~.” |  |
| **Executive summary**Pg 8, lines 27-28 |  | Regarding the sentence “it is the investigators’ responsibility to collect for the study only the sets of data that are strictly needed to provide valid results”, this would be only correct when additional data collection is needed specifically for the study purpose. Otherwise, when only the data collected routinely in the registry are used (then considered as secondary use of data), this cannot be assured as the collection takes place before the study is designed and initiated.*Proposed change*: Specify that the investigators’ responsibility as written only applies to the case when additional data collection is needed. |  |
| **Executive summary**Pg 8, line 28 |  | The term “valid results” is not clear and should be further explained. Would this be linked to robust outcomes? Adequate assessment of confounders? Appropriate statistical analysis?*Proposed change*: Further clarify what is meant here by “valid results”.Also, the term “strictly” is quite vague. *Proposed change:* … that are **defined data fields** ~~strictly~~ needed to provide valid results. |  |
| **Executive summary**Pg 8, lines 32-33 |  | *Data collection*: it is the investigators’ responsibility to collect for the study only the sets of data that are strictly needed to provide valid results. It is also their responsibility to collect all the data needed for this purpose including available data on potential confounders. Data to be collected include those needed for sensitivity analyses as outlined in the study protocol and statistical analysis plan (SAP). The legal status of a study with additional data collection should also be considered. Additional data collection may turn it into an **interventional study** and its relation to current clinical practice needs to be detailed in the study protocol.Clarification and not too narrow a wording would be desirable concerning interventional vs. non-interventional studies**Proposed Change:** Please provide examples to illustrate **specific methods that would be permitted as non-interventional.** |  |
| **Executive summary**Pg 8, lines 34-41 |  | Where data are sourced entirely or largely from EHRs, it is unclear whether the validation should check agreement between the EHR and the registry or validation of the EHR itself. If the EHR is the ‘valid source’ it is unclear when EHR alone would suffice for data source i.e. either as part of an ongoing registry data collection or to obviate the need for new data collection.*Proposed change*: Data quality measures also include quality checks at data entry and monitoring of patient follow-up. **These quality checking standards may not be all applicable or necessary to registries using routine EHRs.**  |  |
| **Executive summary**Pg 8, line 37 |  | Source data verification and periodic auditing *should* be conducted, rather than *may need* to be conducted. |  |
| **Executive summary**Pg 8, lines 38-40 |  | The minimum of 10% of randomly selected patients registered in individual study centres may need to be adapted. Indeed, 10% could represent a quite low number of patients in case of rare diseases, and at the opposite a large number in case of very frequent diseases. In addition, the way the sentence is formulated may be confusing as it refers to individual study centres. Should it refer instead to individual registry centres (meaning outside to any study purpose), or to the registry itself, as there maybe centres that will include many patients and others far less patients where it would be then challenging to do a relevant data source verification on 10% of the registered patients.*Proposed change*: To allow some flexibility here in the **minimum % of patients to be controlled**, and precise if this applies to the registry level, or to each centre participating to a registry. |  |
| **Executive summary**Pg 8 lines 40-41 |  | Quality checks and monitoring should happen in both directions (data “collector” and data owner). For example, if aggregation of data is done at collection level, data owner to validate aggregation level and results to ensure interpretation of variables are correct (e.g. sequence of lab data and results to ascertain an event). |  |
| **Executive summary**Pg 8, line 44-45-46 |  | This is not clear what is meant by “**selected**” in the sentence “they are selected observational cohorts”, especially when selection bias should be avoided as specified in the “patient population” section on page 6.*Proposed change*: Replace “selected” by “**observational cohorts defined** on a specific characteristic”.Profession societies such as ISPR and ISPE has published some good reach practices in this area. It is good to cite these research practices. Here is one example citation by ISPOR-ISPE task force. Berger ML, Sox H, Willke RJ. Good Practices for Real‐World Data Studies of Treatment and/or Comparative Effectiveness: Recommendations from the Joint ISPOR‐ISPE Special Task Force on Real‐World Evidence in Health Care Decision Making. Value Health 2017; (8): 1003-1008. |  |
| **Executive summary**Pg 9, line 12 |  | As important as reporting of study results, is the disclosure of protocol and report on public registers, for registry studies (e.g. PAS register) – importance of transparency. |  |
| **Executive summary**Pg 9, lines 18-19 |  | It is not clear what the right of the MAH is if the registry lead investigator objects/declines to publish the results of the PASS (i.e. primary endpoints for medication under study). This is particularly a problem for MAHs who are committed to publishing research results.*Proposed change*: In the situation where the lead investigator declines to publish the PASS results, **the MAH should have the right to independently publish the PASS** results in due time based on results described in the final study report. |  |
| **Executive summary**Pg 9, line 23 |  | Regarding the “qualification procedures of registries”, there is to our knowledge currently no clear general procedure defined and available to pharmaceutical companies that could be used to ensure the validity of the registries.*Proposed change*: If there is an existing document that the EMA would like the pharmaceutical companies (and registry coordinators) to follow, it should be referenced here. For example, would the Qualification Opinion route be suitable? |  |
| **Section 1 – INTRODUCTION** |
| **1. Introduction**Pg 10, lines 15-17 |  | Industry may also set up disease registries (e.g. to address a post-licensure commitment for a new drug, when rare disease, etc…). |  |
| **Section 2 – OBJECTIVE** |
| **2. Objective**Pg 11, line 3 |  | Reference to “regulatory perspective” is unclear.*Proposed change*: It focuses on important principles ~~from a regulatory perspective~~ **relevant to regulatory decision making, and as appropriate, HTA decision making.** |  |
| **2. Objective**Pg 11, line 20 |  | Also include the ADVANCE Code of Conduct which can apply beyond vaccines.<https://www.ncbi.nlm.nih.gov/pubmed/28285984> |  |
| **Section 3 – CORE CONCEPTS** |
| **Pg. 12, line 4** |  | The definition of patient disease registry is missing here, it would be helpful to delete “patient registry” and replace it with definition of patient disease registry.  |  |
| **Core concepts**Pg 12, line 6 |  | We would **add to the definition of ‘*Patient registry*’: “organised system [...] to collect uniform data on a patient population [...], and which is followed over time at individual patient level** |  |
| **Core concepts**Pg 12, lines 15-223 – 31 (**3.1 Definitions)** |  | *Referred to in Executive Summary, Page 8 (lines 14-26, 18-20), Page 37 (lines 20-23)*“**Incident (prevalent) patient**” is not an appropriate epidemiological term. Indeed, the existing definitions lack clarity as a patient could be considered as prevalent in the context of a registry after which a new registry study is initiated and that patient is then considered to be incident in the context of that study. Consider alternative term for these definitions e.g. newly diagnosed. If retaining the proposed definitions then it should be clear that they are specific to a registry rather than a registry study. |  |
| **Page 12, Lines 2-31** |  | The following additions definitions should be included and/or cross references to GVP Module VI definitions:* Safety reporting definitions (adverse event, adverse reaction, adverse event of special interest)
* Differences between AE or treatment-related events and other (medical) outcomes.
 |  |
| Page 12, Line 17 |  | Comment: The definition of **prevalent patient** includes a first visit; however, depending on how long the registry is willing to accept prevalent use, the visit may be later than the first visit. For example, some US registries allow patients to be prevalent drug users for up to one year (so patient could enroll at second visit). This may be the case with European biologic registers. It may be helpful to consider using the term index date to qualify when the event occurs. Then if index date=registry enrollment, you have an incident patient. If index date < registry enrollment you have a prevalent patient.Proposed change (if any): Line 18…and who enter the registry based on their first visit to a clinician participating in the registry (or as defined by the registry)…. |  |
| Page 12, Line 20 |  | Comment: The date of entry is important for all patients, not just prevalent patients. It is important to be able to differentiate between incident and prevalent patients.Proposed change (if any):Line 16:add: For incident patients, it is important to ensure date of entry in the registry is properly recorded. |  |
| Page 12, Line 28 |  | The reference to population registries includes linkage. Is this unnecessarily narrow? Is the point here that data collected routinely by a government is out of scope? |  |
| **Page 12, line 31** |  | Does the EMA support the use of population registries for providing evidence to support product related assessments of benefits and risks of medicines? |  |
| **Core concepts**3.2Pg 12, lines 36-37 |  | *See also Pg 6, lines 10-11, Pg 10, lines 15-17*As in the executive summary, please add the **pharmaceutical industry (or MAAs/MAHs) in the list of possible registry “coordinator”.** The role of pharmaceutical companies mentioned in 5.8.1.2  foresees only a support of registries Pharmaceutical companies can also be the owner and/or coordinator of registries or disease registries. Companies can be involved at all stages: in the sponsorship, design, execution and analysis of registries. This is particularly noted in rare diseases. For example, the REACH registry is a disease registry sponsored by industry, addressing atherothrombosis and including ~68,000 patients from 44 countries.Should also **expand patients to patient organisations** as this is more consistent with actual practice.It may also be helpful to make clear that registries are **established for different objectives and to set these out.****Proposed change:** Clarify the various roles that pharmaceutical companies can play in support of a registry (sponsorship, design, management analysis) and that a registry may also be created/owned by MAAs/MAHs.*Proposed change*: Patient disease registries are often created by public organisations such as academia or medical research associations of health care professionals or patient **organisations**. **They may also be created by marketing authorisation holders/medicines developers**. |  |
| Page 13, Line 7 |  | Comment: Suggest clarifying what is meant by “controlled designs.” Does control imply an intervention or simply a comparative study? |  |
| **Core concepts**Pg 13, line 20 and pg 14 Table 1 |  | “A patient registry is a data collection system” with a purpose. When it is built is should already consider anticipated research questions which can be analysed later in registry studies.*Proposed change: “A registry is a e data collection system which can sometimes be multi-purpose.”* |  |
| Page 13, Line 22 |  | Comment: Is it worth noting who can initiate, manage, and finance the registry itself (i.e., the data collection system).  As this may be different from who is doing those respective activities for the registry study. |  |
| **Core concepts**Pg 13, line 25 |  | Timelines of a disease registry should not only driven by routine data analyses but also by specific, **anticipated registry study data analyses, which prompted the registry.** |  |
| **Core concepts**Pg 13, line 39 |  | In some cases, the routinely performed analyses, as defined in the routine analytical plan, are descriptive analyses only and do not test a hypothesis as may be the case in registry-based studies or for *ad-hoc* analyses. This will impact the analytical methods used at the registry level or the study level and as such should be referenced in the text.*Proposed change*: *5) Analysis plan*: data analysis in a registry is generally performed at intervals based on patient accrual or time schedule and it follows a routine analytical plan with additional ad-hoc analyses performed by the registry coordinator or registry participants. **In some cases, the routinely performed analyses, as defined in the routine analytical plan, are descriptive analyses only and do not test a hypothesis as may be the case in registry-based studies or for *ad-hoc* analyses.**  |  |
| **Core concepts**Pg 14, lines 1 - 6 |  | *See also Pg 7, lines 12-14*This needs to be clarified, as not all registries set up by a MAH will meet the GVP Module VIII criteria for being a PASS. The collection of safety data alone does not mean that a registry is a PASS, unless the ‘main aim’ of the registry is related to safety (or one of the other criteria established in GVP Module VIII.) Managing a disease registry as a PASS, where it does not meet the criteria for a PASS, would result in over reporting i.e. reporting of all AEs. This may have a significant impact for processing data within registries*Proposed change*: Please clarify the statement to ensure that PASS are correctly identified. |  |
| **Core concepts**Pg 14, line 13 and Pg 15, line 1 (Table 1) |  | *See also comment on Pg 9, lines 4-6,*As raised in the executive summary, within the scope of this discussion paper, studies are based on registry/-ies, either entirely or partially. As such, no study should be considered as being based on primary data collection only.*Proposed change*: For registry studies, requirements to MAH for suspected adverse reactions differ between studies with a design based on ~~primary or~~ secondary data collection **only or, if additional data are to be collected, a combination of secondary and primary data collection**.  |  |
| Page 14, Table 1Data collection |  | Comment: The data collection for the registry study is restricted by the data available in the registry system, not only by what is needed for the research question. Our experience has been that this restriction by the registry system is not trivial.Proposed change (if any): Restricted to **the data available in the registry and to** what is needed by the research question… |  |
| Page 14, Table 1Analysis plan |  | Comment: The registry system should also have a separate analysis plan to detail planned analyses.Proposed change (if any): Under heading RegistrySeparate analysis plan to detail planned analyses |  |
| **Core concepts**Pg 14, lines 20-21 |  | **Reference of registries as “interventional” and “non-interventional”.**Change “interventional” to “low-interventional”. Interventional usually implies randomization and protocol assignment to treatment (or exposure of interest). The examples provided (additional visits or blood tests) are **more consistent with a *low-*interventional observational study**.Currently where an intervention beyond normal clinical practice is required, this is regulated as a clinical trial.Proposed change: Non-interventional. **Note that when the Clinical Trials regulation becomes applicable, certain registry studies might be classed as low-intervention clinical trials.** |  |
| **Core concepts**Pg 14, Table 1, lines 24 |  | In the Table 1 comparison of differences between registries and registry studies, some aspects are missing:* primary data collection versus secondary use of data.
* many registries (eg. Pregnancy registries) have objectives and also have an analysis plan for the periodic data analysis.
* Confusion over interventional and non-interventional regulatory status

*Proposed change*: We suggest two lines for “Analysis objective” and “Analysis plan”:Analysis objective (Registry): Summary of the key registry content and qualityAnalysis plan (Registry): standard analysis plan for regular / routine analysesAnalysis objective (Registry study): Research question as per study protocolAnalysis plan (Registry study): statistical analysis plan separate from the study protocolAlso, please clarify in what kind of scenarios would additional ad-hoc analysis be warranted and who would conduct them? (e.g. new safety signal, analysis by regulator) |  |
| **Core concepts**Pg 14, lines 19-23 and Pg 15, line 1\*\* |  | *See also comment on Pg 24, lies 10-14, Pg 37, lines 39-4*Regulatory status. by nature, a registry is observational, i.e. non-interventional; on the other hand, a registry study may be non-interventional (if data collection is restricted to secondary use of already existing data- see chapter 6.5.) or interventional, for example if additional visits or diagnostic tests are required to validate a diagnostic test in patients identified through the registry, or if additional treatment is given.As indicated in this section, a registry collects data and can be used a data source and is not per se a type of clinical study. Reference to a registry as being “non-interventional” infers that it is indeed a clinical study. This reference should also be amended in the table, as shown below. It is also suggested that requirements for informed consent be included in the table.For the registry study, it would be helpful to have the definition as given on page 37 (ln 17-27)*Lines 19-23**8) Regulatory status*: **a registry does not have a defined regulatory status as it is not a study.** ~~by nature, a registry is observational, i.e. non-interventional;~~ on the other hand, a registry study may **regulated as** ~~be~~ non-interventional **study** (if data collection is restricted to secondary use of already existing data- see chapter 6.5.)**.** or **Alternatively, it may be regulated as** **an** interventional **clinical trial**, for example if additional visits or diagnostic tests are required to validate a diagnostic test in patients identified through the registry, or if additional treatment is given.*Under column: Registry* Regulatory status: ~~Non-interventional~~ **Not applicable**Besides these two “pure forms” there exist registers with defined timeline, predefined number of patients etc. but no research question or hypothesis. This form is a feasible, in specific cases useful form of registries. |  |
| **Pg 14 Table 1** |  | Under “Collection and reporting of suspected adverse reactions”, the sentence “ Any active data collection with involvement of a MAH must follow the regulatory framework for PASS. This is misleading, as not all registries set up by a MAH will meet the GVP Module VIII criteria for being a PASS. The collection of safety data alone does not mean that a registry is a PASS, unless the ‘main aim’ of the registry is related to safety (or one of the other criteria established in GVP Module VIII.) |  |
| **Section 4 USE OF PATIENT DISEASE REGISTRIES IN MEDICINES REGULATION** |
| **Use of patient disease registries in medicines regulation****Pg 16, line 13** |  | “… regulatory authorities may also request MAHs to collect additional data on the utilisation, benefits and risks of their medicines pursuant to a legal obligation (since June 2012) or in the context of the RMP, either as a PASS or as a post-authorisation efficacy study (PAES). Both may be non-interventional, i.e. they may be based on observational data  (20) (22).”  Proposed change, in order to avoid misunderstanding: “Both may be interventional or non-interventional, i.e. ..  |  |
| **Use of patient disease registries in medicines regulation**Pg 16, lines 16-17 |  | “…with a special population affected by a disease, such as older people”. The “such as” example may not be appropriate as the older people are not always affected by a disease.*Proposed change:* Maybe find another example or complement such as “older people at risk of developing Alzheimer’s Disease”. |  |
| **Use of patient disease registries in medicines regulation**Pg 16, lines 17-20 |  | **There may be a confusion here between the registry as a source of data and the registry-based studies**. Indeed, patient disease registries are mostly based on data collection from the patients by the registry participants. Then, they would not be considered as secondary data collection. In contrast, studies using data from registries would be considered as studies based on secondary data use, as the data would have been already collected in the registries, when used for the specific study purpose and other scenarios.*Proposed change:* Please clarify when referring to registries or registry studies. Consider deleting “or secondary data collection (…)”. |  |
| **Use of patient disease registries in medicines regulation**Pg 16, lines 17-22 |  |  “The registry can be used as a source of patients based on either primary data collection (where the events of interest for the study are collected directly from patients as they come to the attention of the investigator) or secondary data collection (analogously to the use of electronic healthcare records)”It needs to be carefully evaluated if a systematic secondary data analyses referring to own products is possible, especially if primary data was not independently collected. *Proposed Change*: It should also be described that a registry could be used to embed a randomized study within the registry, i.e. to identify patients meeting inclusion/exclusion criteria and follow the study outcomes in the registry. “For this purpose, registry data can be enriched with additional information from linkage to existing databases such as national cancer registries, prescription databases or mortality records. “It is an exciting approach to enrich registry data with additional information.**Proposed Change:**It would be helpful to have guidance on methods and clear controls/criteria on how these data could be access and analysed.  |  |
| **Use of patient disease registries in medicines regulation**Pg 16, lines 23-25 |  | “*For common diseases, exposures or conditions, alternative data sources providing access to large numbers of patients, such as electronic healthcare databases of medical records or claims data, should be rather considered”.* This statement is not entirely correct and cannot be generalized. Relevant data is often missing from EMR/EHR sources. Data sources providing access to large numbers of patients for common diseases, exposures or conditions have certain advantages, but may not contain the type of rich clinical data that is often available within disease registries (e.g., patient-reported outcomes). There are examples of registries for more common diseases (e.g. clozapine registry, psoriasis registry [BADBIR]) that are very valuable. The selection of electronic healthcare databases over registries **depends on availability of key variables** rather than sample size.These large data sources may be better characterized as **complementary to, rather than a replacement for,** disease registries. The current text may suggest that the latter is being recommended.EHR databases are often small and claims data often lack important clinical information. In some cases, the registry sizes (even for diseases that are not rare) are small. A statement that makes these considerations may be better.*Proposed changes*: Page 16, Lines 16-17 “They are particularly useful when dealing with a rare disease or with a special population affected by a disease, such as older people *{at risk of developing e.g. Alzheimer’s Disease}* **but may also be of benefit for more common conditions**.”Page 16, Lines 24-25: Suggest changing ‘…should be rather considered’ to ‘…should also be considered in addition to patient disease registries’Add line 25However not all electronic healthcare databases are large and claims databases can lack clinical information |  |
| **Use of patient disease registries in medicines regulation**Pg 16, lines 26-40 |  | The examples given here as possible studies to be performed in registries are only linked to the specific context of the medicines regulation and the post-marketing setting (and for this purpose, the case of “historical control data that could be used for comparative purposes” would not really fit the post-authorisation setting). Registry data could also be used in the pre-marketing setting as described at the end of the first paragraph of this page. In the PASS bullet, use of registries for historical controls may be applicable here as well.*Proposed change*: to remind the specific context here (**mostly post-authorisation setting**) and adapt the text accordingly (**include studies on historical control data as a possible pre-authorisation setting)**.If however the intent is to focus on scenarios where studies can be mandated, then this intent in the range of possible studies should be set out. |  |
| **Use of patient disease registries in medicines regulation**Pg 16, line 28 |  | We suggest to add “...useful for the following registry studies:” to especially make clear that DUS, PAES and PASS are examples for registry studies but not for registries. |  |
| **Use of patient disease registries in medicines regulation**Pg 16, line 39 |  | We suggest to add “Registries **and registry studies** may be particularly valuable when examining the safety of medicinal product used for an orphan disease.” To avoid confusion |  |
| **Use of patient disease registries in medicines regulation**Pg 16, line 44 |  | This is an interesting approach but there should be caution on “generalizing” RCT data unless the RCT is a cohort selected from the registry. |  |
| **Use of patient disease registries in medicines regulation**Pg 16, line 35-40 |  | Please align the section about PASS with EU GVP Module 8 “post-authorisation drug safety studies (PASS), to collect safety data on adverse events using standardised data collection tools and amplify a safety signal, particularly for rare outcomes, to assess the incidence of important identified and potential risks, to compare the risk of some adverse events between relevant exposure groups or to assess the effectiveness of risk minimisation measures. Registries may be particularly valuable when examining the safety of medicinal product used for an orphan disease.”*Proposed change:* post-authorisation drug safety studies (PASS), to quantify potential or identified risks, to evaluate the risks of a medicinal product used in a patient population for which safety information is limited or missing, to evaluate the risks of a medicinal product after long-term use, to provide evidence about the absence of risks, to assess patterns of drug utilisation that add knowledge regarding the safety of the medicinal product or the effectiveness of a risk management measure and to measure the effectiveness of a risk management measures. |  |
| **Section 5 GOOD REGISTRY PRACTICE** |
| **Good registry practice****5.1 General considerations**Page 18, line 3 |  | Some guidelines describe different designs for a registry, including the traditional case series, cohort, case-control or case-cohort designs.Consider including references for “some guidelines” in order to make it clear which guidelines are implied. |  |
| **Good registry practice****5.1 General considerations**Page 18, lines 3-8 |  | Some guidelines describe different designs for a registry, including the traditional case series, cohort, case-control or case-cohort designs. This reflects the misrepresentation of a registry as a study rather than a data collection system. This reflection paper establishes that a registry, as a system that collects data on patients followed-up over time, inherently leads to the creation of a cohort of patients that may be secondarily used to investigate a research question through different study designs. Therefore, registry designs will not be discussed.it is important to understand **the view concerning the concept of sub-studies** **such as registry-based studies**. it was mentioned that patients of a registry may be secondarily used to investigate research questions in a separate study, however a possible design of such studies was not further discussed. Use of “registry designs” here perpetuates the misrepresentation of a registry as a study rather than data collection system.   It is important to recognise that a registry can “feed” multiple registry studies.**Proposed Change:** “Registry **study** designs will not be discussed”.To understand the concept and study design of registry-based studies it would be appreciated if the Agency could provide further information about possible designs of studies where patients of a registry may be secondarily used to investigate research questions in separate studies. |  |
| **Good registry practice**5.1 General considerationsPage 18, lines 10-15  |  | A clearly defined purpose and objective for developing a patient registry at the beginning can contribute to minimizing bias.This refers to “i.e. the situation where enrolment is influenced by patient characteristics that may affect the validity of the analyses”. However, the analyses may be “valid” (in that they accurately address the question as it is defined), but they may be addressing the wrong population. (if, for example, only patients with more favorable prognoses are included in the registry.)It should be clarified whether the text is intended to address the possibility that the outcome assessment may not be representative of the outcomes in the target population of interest Proposed changes:1. *Addition:* Defined purpose and objective for developing a patient registry
2. *Clarification*: Is the text intended to address the possibility that the outcome assessment may not be representative of the outcomes in the target population of interest?

  |  |
| **Good registry practice**5.1 General considerationsPage 18, lines 11- 13 |  | We suggest to add “Great care should be exercised to ensure a comprehensive enrolment of patients and avoid selection bias, i.e. the situation where enrolment is influenced by patient characteristics that may affect the validity of the **registry** analyses.” This is to emphasize the importance of good quality registries for the registry studies. |  |
| **Good registry practice**5.1 General considerationsPage 18, line 20 |  | We suggest to add...or another purpose to support registry studies including a “list of anticipated registry studies with its study objectives.” This might help adding clarity how important up-front thinking about future studies is so that an appropriate set of core data elements can be collected. |  |
| **Good registry practice**5.1 General considerationsPage 18, line 21 |  | We would suggest to add: *“…*needs in-depth consideration and consultation **of all relevant stakeholders, including** specialists and patients concerned”. Since e.g. coordinators of other registries and KOLs should be part of this discussion as well. |  |
| **Good registry practice**5.1 General considerationsPage 18, line 30 |  | We suggest to add “...of effect **modifiers in registry studies**.” This might help adding clarity on how important follow-up information is for registry studies. |  |
| **Good registry practice**5.1 General considerationsPage 18, lines 37-39 |  | Would this mean to compare all eligible patients in a region to those in the registry of the region? Our understanding of the above-mentioned is that the goal would be to include all eligible patients of a region, hence the target population would be equal to the eligible population. Where would the information on patients that would be eligible but not willing to participate in the registry in the first place come from? How could they reliably be identified? |  |
| **Good registry practice**5.2 Patient populationPg 18, lines 26-27, 30 |  | An exhaustive enrolment may not always be the ultimate goal for a registry, especially in the case of common diseases (as listed on the same page, lines 31-33), and also as some patients may refuse to sign the informed consent form and be enrolled. We believe it only possible for population registries, which are out of scope for this Paper, or at least near complete coverage of the population of interest.What is more important in case of common diseases should be to ensure a certain degree of representativeness for the enrolled patient population. Documentation of the number of patients who were asked to join the registry and those were included could provide evidence of the representativeness.This **refers to assessing registry representativeness** using data from electronic records- as in comment for page 8, lines 34-41 above, it is assumed that the EHR is considered the ‘valid’ source but there is little information about what the acceptable variation would be in terms of representativeness (or if statistical adjustment is proposed).Terms such as “exhaustive” and “complete” are important, in theory, but very hard in practice to realize. *Proposed changes*: 1. Minimize the statement regarding the exhaustive enrolment and ensure that this is adapted to the frequency of the disease of interest. Consider softening the language which could lead into the text on line 34 which considers “non-exhaustive” sampling schema.
2. *To add*: …These may not all be applicable or necessary for population-based registries (e.g. those built from existing EHR systems).
 |  |
| **Good registry practice**5.2 Patient populationLines 35-36 |  | Comment: While a good suggestion for attempting to allow for exhaustive enrolment, this will limit generalizability and this type of selection may lead to a bias as well.  Consider noting this.Proposed change (if any):…or other criteria. **However, this will limit generalizability and this type of selection may lead to a bias** |  |
| **Good registry practice**5.2 Patient populationPg 19, line 3 |  | **The benchmark of the actual registry population in comparison with other data sources should not guide the recruitment strategy for a *specific study*, but for the registry itself.** For registry-based studies, the recruitment strategy will be based on the registry data itself. But the benchmark with another source could guide the strategy regarding the additional data collection to be done for some specific studies. Comparison of registry data with health care records may not be possible in certain countries due to national data protection laws.*Proposed change*: Replace “for a specific***study***” by “for a specific **registry**”.  |  |
| **Good registry practice**5.2 Patient population |  | While an ideal situation would strive to ensure exhaustive enrolment; it’s more realistic to frame this as methods which maximize chances for exhaustive enrolment.  Another perspective is that exhaustive enrollment is not always needed. Sampling methods can be considered to ensure efficient estimation of target parameters.Similarly, complete f/u is ideal but not realistic. The goal is to create a system that best supports complete f/u (i.e., what steps are taken to minimize loss to f/u). |  |
| **Good registry practice**5.2 Patient populationPg 19, line 14-37 |  | The selection of electronic health databases over registries depends on availability of key variables rather than sample size, the collection of information about patients who did not consent is limited per data privacy – see also similar comment on Pg 16, lines 23-25 of section “Use of patient disease registries in medicines regulation”. |  |
| **Good registry practice**5.3 Time elements |  | Some guidance would be welcome to bring together the 5.3. Time elements and the 5.4 Core data elements to clearly show the link between both data sets. Perhaps consider integrating the two sections to achieve this. Otherwise, there is the risk to generate inconsistent information (e.g. gathering date of symptoms but no symptom). |  |
| **Good registry practice**5.3 Time elementsPg 19, line 6 |  | Important events may be seen as only related to safety or effectiveness events, and may be seen as too restrictive when dates should be linked to important data or variables to be collected.Proposed change: Replace the term “**events**” by “**data or variables**”. |  |
| **Good registry practice**5.3 Time elementsPg 19, lines 5-19Also Table 2 |  | Additional clarity would be beneficial on which core events from *prior* to the registry entry date are essential for capture in the registry. For some components (e.g., date of first symptoms) it may be assumed that the event typically occurred prior to registry enrolment. However, for recommended components such as treatment start/stop and investigation dates, it is not clear whether this applies to occurrences prior to initial registry enrolment. This is particularly important for registries which include prevalent disease cases, as knowledge of events prior to registry enrolment may significantly impact safety, effectiveness, and other outcome analyses. *Proposed change*: **Clarification needed whether certain core dates refer to both pre- and post-registry enrolment events** |  |
| **Good registry practice**5.3 Time elementsPg 19, line 18 |  | When specifying which time elements should apply across different registries of patients with the same disease, it would be beneficial to consider a central repository e.g. at EMAs homepage where all stakeholders can see expected time elements per disease registry as they become available and then introduce/refer to such a repository here. |  |
| **Good registry practice**5.3 Time elementsPg 19, line ca. 30 |  | The table is missing the steps to be clearly documented to demonstrate representativeness.*Proposed change*: Add this point between lines 30 and 31. |  |
| **Good registry practice**5.3 Time elementsPg 19, line 19Table 2Pg 19, table 2 line 30Pg 19, table 2 line 31-32Pg 19, table 2 line 35-36Pg 20, table 2 line 12 |  | The table specifies that the exact date of birth be collected. It needs to be confirmed that this is allowed to be collected for registries according to GDPR and national regulations. We believe that an ICF signature would allow for the collection of birth date. Patients should be able to opt out of providing birth date and to provide age instead, e.g.One basic obligation for all data controllers is to ensure that they process only the personal data that is relevant and necessary for the achievement of their lawful, fair and legitimate purposes. As a result, any personal identifiers that are collected and which could lead to identifying a data subject, directly or indirectly (or in conjunction with other data) shall be evaluated under the principle of data minimisation. If less intrusive means are available, it shall be always preferred. The above apply to the info mentioned in the table, which could lead to identification of data subjects (full date of birth or death mainly). It is always a matter of debate whether a range (e.g. of years) or a year of birth or death (without full date details) would suffice to accomplish the task sought in an accurate and effective manner. This shall be the preferable and default position. Consider adding index date and a flag for incident/prevalent patient. Also, some dates are specified as exact dates and others are not. What is the reason for this?Under ‘disease dates’: an intermediary between symptom onset and actual diagnosis could be added: date of referral to specialist (particularly relevant for diseases with long/complex natural history / time lag between first symptoms and final Dx). Under ‘Other event dates’: suggest including medical conditions not directly related to treatment – AE will need to be defined. Under ‘other events dates’: how about other relevant medical history? (provided link to such information is feasible)* Under ‘other event dates’: what is meant by “significant@? Should this say “serious”? Are these synonymous with non-protocol defined AEs, whereas the AESI are protocol defined? Suggest to clarify the difference between AESI dates and other events dates.
* ’Date(s) of cure or significant improvement(s)’ and ‘Date of relapse’ may either not be applicable or be difficult to define and capture for certain diseases. Specifically, disease activity and patient status for many chronic diseases may be tracked over time via several metrics, including laboratory tests, clinical examination, patient report, etc, without an easily defined ‘significant improvement’ or ‘relapse’ date.

*Proposed changes*:Combining sections 5.3 and 5.4, or at least adding elements (bullets) on Pg 21 to Table 2 to have a table encompassing all core elements.To add: “…Steps to be clearly documented to demonstrate representativeness”1. To add: Under ‘disease dates’: an intermediary between symptom onset and actual diagnosis: date of referral to specialist (particularly relevant for diseases with long/complex natural history / time lag between first symptoms and final Dx)
2. To add or modify: disease dates to more broadly reflect change in clinical disease activity or disease status versus a focus on ‘significant improvement’ and/or ‘relapse’
3. Replace “**AE**” by “**medical conditions not directly related to treatment**”
4. To add: “other relevant medical history” (provided link to such information is feasible)

Date of birth: Please recommend clear standards (dd/mm/yyy) to allow pooling later on (relevant for the whole core list as one example)Treatment dates: please add dosage because different dosages will start/stop at different times.Relevant co-therapy: please define relevant; in order to detect potentially new safety signals relevant co-therapy should be considered broadIt needs to be confirmed that this is allowed to be collected for registries according to GDPR and national regulations. |   |
| **Good registry practice**5.4 Core data elementsPg 20, line 12-13 |  | “Core data elements should be defined with clinicians and experts concerned by the disease as well as representatives of patients and end-users of registry information.”*Proposed change:*To add: and should be in line with existing clinical trials guidelines for clinically relevant outcomes (e.g. <https://www.ema.europa.eu/en/human-regulatory/research-development/scientific-guidelines/clinical-efficacy-safety-guidelines>The use of validated outcomes or laboratory tests should be prioritized and harmonized through different registries (e.g. Quality of life questionnaire SF-36. Core data should comprise baseline characteristics which are in line with current practice in clinical trials in order to provide the basis for indirect matched comparisons of real world data with data from single arm studies, in case randomized clinical trials are not possible.  |  |
| 5.4 Core data elementsPg 20, lines 12-13 |  | Core data elements should consider multiple geographical area, multiregional registries should harmonize the data elements and ensure variation in local SOCs are not barriers for data collection.Core data elements need to furthermore define composite scores as these may vary between physicians regions and countriesThis should also consider the dynamic of a registry; meaning e.g. diagnostics, SoC, treatment list and regular procedures evolve overtime. The need of core data elements updates should be clearly reassessed on a regular basis during the conduct of a registry and especially for long-term and open-ended (e.g. more than 10 years) registries. **Proposed change:** Core data elements and their composite scores should be defined with clinicians and experts concerned by the disease, as well as representatives of patients and end-users of registry information. The data elements should be defined and aligned with local routine clinical practice, prescribed dose should be mentioned for therapies and geographical coverage should be taken into consideration together with lifestyle factors **(such as smoking, alcohol consumption, physical activity, nutrition…)** as consistently as possible. The data elements should be revised during registry life to ensure it is still aligned with routine clinical practice, a critical review of missing key data elements should be done on a regular basis. |  |
| **Good registry practice**5.4 Core data elementsPg 20, line 18 |  | Examples of core data elements also exist in infectious disease surveillance systems, which can be related to registries.*Proposed change:*To add: …. core data elements “**exist in infectious disease surveillance systems**” and agreed for disease… |  |
| **Good registry practice**5.4 Core data elementsPg 20, line 25 |  | Typographical error*Proposed change:* Replace “**date**” by “**data**” in the sentence “The adoption of a common set of core date elements across registries does not exclude…”. |  |
| **Good registry practice**5.4 Core data elementsPg 20, Line 26 – 28Pg 21, line 1 – 18Pg 21, line 1 - 3Pg 21, line 4-5 Pg 21, Line 7Pg 21, lines 11-13Lines 11-12Line 17 |  | The “should have” denomination is a bit misleading as for the core data elements, the term “should” is also used (see page 20, line 27). The recommendation would be to make it clearer which core data must be collected and which data it is not obligatory to collect.The text initially refers to “crucial elements,” then mentions “should have” elements and then back to a more detailed listing of “crucial” data elements. This seems potentially confusing. Suggest changing the order to: high-level mention of crucial elements, then details of crucial elements, then high-level mention of “should have” elements. Consideration should also be given to providing some examples of “should have” elements. *Proposed changes*:1. “Crucial” data elements are those which should be collected in all registries and on which greater amounts of resources should be allocated to ensure completeness, standardisation, data quality and verification of the information. Consider mentioning endpoint adjudication by trained professionals specifically.
2. “In the perspective of ~~regulatory~~ registry studies, “**crucial” data elements** should include the following information:-Patient data: …….
* ....

-Pregnancy: date of start of pregnancy, pregnancy outcome (spontaneous abortion, live birth, etc.).1. **“Should have” data elements** are those considered of interest and useful for some stakeholders, but not essential to all. Examples include: **x, y, z**
2. To add: Besides “crucial” vs. “should have”, there is also a need for “**standardised data format elements**” and “**variable definitions**” to allow merging between registries or different data sources, depending on the disease, product, intention of the registry.
3. To delete: crucial data elements which may be too disease-specific
4. To add: race/ethnicity information for patient data
5. To add: diagnosis onset date
6. Clarification: “*burden of disease*” – what data are expected?
7. To add: Include in disease-related treatments data beyond drug-related information, and precise for the other therapies that this is about therapies used to treat other diseases than the disease which has triggered the inclusion into the registry.
8. To replace: “**reason for discontinuation**” by “**treatment change rational/reason**” as patient may discontinue, switch or augment with other medications
9. To add: ... “pregnancy outcome (spontaneous abortion, live birth, “**preterm birth, still birth and infant outcomes of major anomalies and hospitalised infections**”, etc.)
 |  |
| **Good registry practice**5.4 Core data elementspage 21line 11-12 |  | **Disease**: list of core data elements to be determined for each disease with core data on diagnosis (date, test, result), grade/severity/burden of disease, important milestones in disease progression and core disease outcomes (e.g. relapse, disabilities, functional status, quality of life measure, cause of death);**Proposed change:**In each disease registry clinically relevant prognostic factors for each disease should be included as core dataAdverse events of special interest and serious suspected adverse reactions: adverse event MedDRA terms, start and end dates, treatment suspected to be associated, seriousness, dechallenge, outcome of adverse event;**Proposed change:**Safety data shouldn't be restricted to AESIS, additionally AE 3-4, 5, SAE and AE leading to discontinuation should be included as core data |  |
| **Good registry practice**5.4 Core data elementsPg 20, line 28 |  | Completeness and standardisation should be part of the data quality and checked via verification of the information.*Proposed change:* Replace “to ensure completeness, standardisation, data quality and verification of the information” with “**to ensure data quality (including data completeness, standardisation and accuracy) thanks to verification of the information**”. |  |
| **Good registry practice**5.4 Core data elementsPg 21, lines 4-5 |  | There is a **question about the need to include the following variables in the list of crucial data elements** (which means mandatory data collection) is not obvious, as is highly dependent of the disease of interest and research questions: height, weight, indicator of socio-economic status (e.g. highest educational level completed), smoking status. These variables could be highly dependent on the disease of interest and research questions. However, these data may have an effect modification in certain indications.*Proposed change*: To review the need for data which may be too disease-specific in f the list of crucial data elements.Besides “crucial” vs. “should have”, there is also a need for standardised data format elements and variable definitions to allow merging between registries or different data sources (again, depending on disease, product, intention of the registry).Suggest adding race/ethnicity information for patient data |  |
| **Good registry practice**5.4 Core data elementsPg 21, line 7 |  | It is not obvious what is covered by “burden of disease”, and then why this is part of the crucial data elements. We note that the burden of disease will depend on the disease.*Proposed change*: Clarify what is expected to be collected regarding the “burden of disease”.For diagnosis, need also to include onset date. |  |
| **Good registry practice**5.4 Core data elementsPg 21, lines 4-19 |  | EMA nicely encourages to use existing data dictionaries. However, the core data elements should then also list the “data codes” (i.e. disease code, test code, treatment code, etc.; as these are derived data it does not increase the burden on data collection.) |  |
| **Good registry practice**5.4 Core data elementsPg 21, lines 13-15 |  | Adverse events of special interest and serious suspected adverse reactions should only be part of the core data elements, if their report is considered routine clinical care. Otherwise, they should only be collected as part of a registry study, hence not changing the interventional status. |  |
| **Good registry practice**5.5 TerminologiesPg 22, Table 3 |  | Many of the recommended international terminologies are already in place for EHR, claims, and they still pose challenges and can benefit from further harmonization, e.g., common data model across. Please note on international terminologies: Reference to the Classification of functioning/disability International Classification of Functioning and Disability (ICF); ECOG is common and used in and outside of studies. ICF has to be licensed. |  |
| **Good registry practice**5.5 TerminologiesPg 22, Table 3 |  | **Rare disorders.** Orphanet managed the definition and inclusion of an additional 5000 rare diseases in the ICD-11, why is the ICD-11 not listed as one of the standards/ cross-references for rare diseases? |  |
| **Good registry practice**5.5 TerminologiesPg 23, lines 5-7 |  | It should be clarified that reference to the use of SNOMED terms, derived from mapping of ICD codes, is acceptable. This is important as SNOMED is standard for OMOP Common Data Model.*Proposed change*: Another clinical terminology **that may be used** is the SNOMED Clinical Terms (CT) terminology. |  |
| **Good registry practice**5.6 Quality managementPg 23, lines 38-41;Pg 24 lines 1-4 |  | “Plan, Do, Check, Act” should preferably be done in a continuous manner all along the lifecycle of the registry. Ownership and responsibility should be clearly defined to enable sustainability of this quality management system |  |
| **Good registry practice**5.6 Quality ManagementPg 24, lines 6-7 |  | There should be transparency on the quality assurance activities carried out in the registry by the coordinator so that potential users of the registry data are able to assure themselves that appropriate quality control is being performed e.g. through quality assurance reports. |  |
| **Good registry practice**5.6.2 Requirements |  | Data quality and appreciation of data quality should always be put in context of the information the HCP participating to the registry network has available (e.g. what is his/her position on the patient pathway? Timeliness is driven by when the physician has been made aware of the event, if he is not the physician centralising all the information of the patients), accuracy and completeness can only be assessed in regard to the information the physicians received (e.g., full lab test results vs, letter from a peer in a different specialty).  |  |
| **Good registry practice**5.6.2 RequirementsPg 24 lines 10 – 40 |  | What is considered acceptable data quality? Are there thresholds defined for the four main components of data quality. When and how will we know that measures to improve data quality need to be implemented?**Minimum required quality standards for registries used by regulators are lacking, and more granular guidance needs to be given with regards to what defines quality data**. Based on registries that have been the subject of an EMA qualification procedure, minimum standards might be proposed e.g. required % completeness of data; data should be medically confirmed (perhaps via site monitoring). |  |
| **Good registry practice**5.6.2 RequirementsPg 24, lines 11-12Lines 10-14 |  | Consistency should not be limited to be over time and across different registries, but also across different centres as there is sometimes the risk that, for example, a same disease to be coded differently between different centres. The standardisation may minimize this risk, but not avoid it completely. Only a clear guidance and trainings can ensure that data are consistently recorded within the different centres.*Proposed change:* To modify as follows “…are consistent over time and across different centres within the same registry and different registries…”There could be a metadata file detailing the meaning of each variable so that all stakeholders are aligned when mapping/entering data.Line 11: *Consistency: the formats and definitions of the* variables and [please add] *metadata related to necessity of fields…* This could be used in the long term to assess the quality of the registry, but we need to factor in the critical vs. the nice to have elements.It can be recognized that data completeness is expected to be low for certain outcomes especially in comparison with RCTs. Thresholds of >90% cited as example in Table 4 can be challenging. This should be taken into account in feasibility assessment and highlight the importance of missing data handling. |  |
| **Good registry practice**5.6.2 RequirementsPg 24, lines 18-19 |  | See previous comments about EHR and data quality, substitution of data from health record, representativeness, and lack of detail on assessing representativeness, adjustments for discrepancies in representativeness and data quality, etc. |  |
| **Good registry practice**5.6.2 RequirementsPg 24, lines 20-21 |  | (Timely recording) Also when there are updatesto data |  |
| **Good registry practice**5.6.2 RequirementsPg 24, lines 30-32 |  | “Verification of clinical diagnoses through data linkage to hospital data or verification of the information on drug exposure in the study population through data linkage to a claims database” is possible for specific registry studies but not for the disease registry as such. This should be included in the section on registry studies.  |  |
| **Good registry practice**5.6.3 Measures to improve qualityPg 25, lines 9, 15-16 |  | ***“Registry coordinators should provide to local registries and centres harmonised definitions and data elements; a support function should be made continuously available at central level; use of a common data collection and reporting software should be considered.”***This is the first time the below roles are mentioned. It might be good add more information about these roles (see also comment under Page 5):* Registry coordinator
* Local Registry coordinator
* Data custodians
* Data manager

*Proposed change*: Add definitions of the above roles here or within the documents. Consider adding the bullet on site-specific statistical programming.Who would be considered the data custodian? Is this term interchangeable with registry coordinator? If so, please align terminology. |  |
| **Good registry practice**5.6.3 Measures to improve qualityPg 25, lines 17-18 |  | Suggest modifying “national databases” to “national or other e.g. EHR or insurance claims, databases.” The latter may not be national.*Proposed change*: If legally and technically possible, a linkage system with other national databases **(or non-national databases such as those for EHR, insurance claims, etc)** to double-check data or extract additional information should be established. |  |
| **Good registry practice**5.6.3 Measures to improve qualityPg 25, lines 19-21 |  | Who should oversee these analyses; individual centres, national sites or registry coordinators? |  |
| **Good registry practice**5.6.3 Measures to improve qualityPg 25, lines 28-31 |  | **Regulatory qualification c**ould be encouraged in order for registries to robustly support public health. |  |
| **Good registry practice**5.6.3 Measures to improve qualityPg 25, lines 34-35 |  | In addition to trainings of data managers and other persons involved in data entry, which is critical to ensure accuracy of the data entry, ensuring that these persons are qualified for this and have a minimum knowledge about the disease captured in the registry cold be very valuable to ensure data quality.*Proposed change*: Consider to add appropriate qualification requirements of personnel involved in data collection process, in addition to trainings. |  |
| **Good registry practice**5.6.3.1 Measures to improve quality**\*\*\*** |  | Section 5.6.2. states “A good balance will need to be found between introducing additional control measures, avoiding redundancy in data collection and keeping a manageable time for data entry, as a cumbersome data entry process may increase the amount of missing data.”Under the same rationale, avoiding the collection of too much/redundant data could also be defined as a “measure to improve data quality” |  |
| **Good registry practice**5.6.4 Indicators |  | Although it may not be possible or desirable to fully define, the reader would benefit from any further clarification of the meaning of terms: “acceptable” data quality, “successful” remedial actions and when exactly their importance may indicate higher/ lower “requirements of data quality”. |  |
| **Good registry practice**5.6.4 IndicatorsPg 27, Table 4Proposed indicators |  | Regarding data accuracy, validation against source data for 10% of registry data is not feasible in all cases and should rather be done on a risk-based approach.Who will be in charge of quality assessments – **registry owner, third party?**Suggest replacing “*Variability across fields*” by “*Variability within and between fields*”Proposed indicators of accuracy – depending on the data source (e.g. claims database) and the amount of data points to be verified, a 10% cross form validation is often not realistic in terms of cost and availability of e.g. electronic charts.Proposed indicators of completeness [ROW 3, COLUMN 3] – although the use of core common data elements is supported, these are still to be defined for many disease areas/outcomes and hence currently their use is not always feasible. It would be too burdensome to both EMA and MAHs to have a scientific advice/qualification meeting for each disease (aside from where a registry is the basis of a study that will be reviewed by PRAC). However, partnerships, between regulatory authorities and companies those with access to data, could be an effective link. Furthermore, companies are not currently mandated to seek such input at other stages of development.Proposed change (if any): Agreed list of data elements and definitions **(optional)** |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 3-4 |  | Should not all disease registries conducted by organizations follow the national requirements for the management of safety data? What type of organisations would be excluded?Does that mean that the CIBMTR, being an umbrella registry, **needs to be following national requirements for all nations data is pooled from or does that responsibility remain with the (national) registries the data comes from?** |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 3-8 |  | How will it be possible to deal with duplication of reports? |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 7 |  | The abbreviation “NCA” is used for the first time in the document and it may not be known to everybody.*Proposed change:* Please use the full term “National Competence Authority” together with “(NCA)” to introduce the acronym. |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 9 – 17 |  | To our understanding, this paragraph addresses registry studies and not good registry practice by having an MAH funded active data collection. Please consider moving it to chapter 6.  |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 9 – 11 |  | Related to see comment on Page 7, Lines 12-14 above. *Proposed change*: Please clarify that this applies when collection and recording of AEs is the main objective of the activity (per GVP Module VIII) i.e.Any active data collection system put in place in a disease registry, fulfilling the PASS definition, and initiated, managed or funded by a MAH must follow the regulatory framework for PASS (20) (21) (see chapter 6.8).  |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 14-16 |  | In the sentence “Relevant safety information should be summarised in the product Periodic Safety Update Report (PSUR) by the MAH and in the registry reports. “ it needs to be clear that it is registry study reports and not registry reports and it should be aligned with GVP module 8 to avoid confusion.Proposed change: “Relevant safety information should be summarised in the product Periodic Safety Update Report (PSUR) by the MAH and in the **registry study reports (i.e. interim safety analyses and final study reports)**.“  |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 9-17 |  | This paragraph is referring to data collected for registry studies according to a protocol that are conducted by a MAH. According to GVP requirements and internal procedures MAH is obliged to explicitly record and report suspected adverse reactions to its medicinal products regardless of the registry/study aim. The definition in its current form could be interpreted that a disease registry initiated by MAH could therefore fulfil PASS requirements, even if it is initiated purely with the aim to study disease epidemiology etc. Furthermore, PASS definition should be only applicable to registry studies rather than registries.Therefore this safety reporting aspect should be addressed in the section on registry studies (chapter 6.8) rather than here, as it is rather confusing to mention protocols in this section.*Proposed change*: Suggest deleting lines 12 – 17 and inserting this information in chapter 6.8. Make it clear that PASS designation may apply to studies only, but not to registry itself as a form of data collectionActive safety data collection by the MAH and differentiation between the requirements for solicited and spontaneous reporting is not always clear cut in increasingly more complex study designs. When safety data are collected only for specific adverse events/outcomes related to a medicinal product, **please clarify** **how safety reporting should be managed for any other adverse events/reactions incidentally identified with the product within the same study** i.e. reported as solicited events, with causality assessed accordingly, or handled as spontaneous reports, where appropriate. |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 32-33 |  | Missing information listed as such in the RMPs should not always be considered as AESIs, as often related to specific patients’ sub-groups not assessed during clinical development.*Proposed change*: Adapt the text accordingly to avoid linking missing information in RMPs to AESIs. |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 36-42 |  | **Aggregate analysis of adverse events**Routine statistical analyses of collected adverse events could be based on standard statistical programming that can be run periodically by registries or upon requests from a regulator or a MAH and agreed with them. Analyses can be descriptive but can also integrate comparisons between different treatment groups (based on therapies, dose, duration, etc.) or categories of patient characteristics. Routine statistical analyses should be defined in advance and described in a SAP. Comparative analyses could require defining comparator exposure groups and controlling for confounding factors, especially confounding by indication. Consideration should also be given to various treatment periods to which patients may be exposed, especially in case of switching between therapies, in order to avoid time-related bias (hence the need for an accurate recording of times of entry into the registry and start and end times of different therapies).**Proposed change: MAHs should be informed and data should be shared with MAHs as well** (registry may not always be MAH registry) and get the chance to discuss results. Standard procedure from regulators on how to deal with missing data/variables and unpublished data seems desirable/inevitable.Further guidance is welcomed on safety reporting in case of multiple MAHs |  |
| **Good registry practice**5.7 Safety Analysis Pg 28, lines 36-37 |  | The audience will benefit from further clarity if the document can include several established methodologies for confounding control, e.g., propensity score, causal inference.Suggest saying that comparative analyses should require defining comparator exposure groups, etc.*Proposed change*: Comparative analyses **should** ~~could~~ require defining comparator exposure groups … |  |
| **Good registry practice**5.8 GovernanceP 29, Line 7 |  | As a general comment, EFPIA’s observations on governance are extensive and fundamental. We believe it would be of **benefit to set up a follow-up expert discussion** on this subject separate from any other stakeholder engagement that EMA may have planned in relation to this process |  |
| **Good registry practice**5.8 Governancepg 29 lines 31-39 |  | The proposed policy has some common elements with privacy notices required under GDPR and should be aligned with the GDPR requirements. It is unclear what is meant by "data ownership" (line 35) clarification would be useful. It is not clear what the rationale is behind the inclusion of the phrase "possibilities for pooling......other data sources" in the scope of the institutional policy. The policy should make clear whether data being shared is personal data in GDPR terms and inform recipients of their obligations under GDPR |  |
| **Good registry practice**5.8 GovernancePg 30, lines 1-8 |  | Is this not a breakdown and specification of the above bullet point “process for managing requests” (page 29 line 40-41)? Could this be somehow marked or otherwise specified what additional/ separate processes are in place? |  |
| **Good registry practice**5.8.1, Governance principles, line 5 |  | The scenario where specific medical societies would have a monopoly of coordinating particular disease registries should be avoided. Data ownership, intellectual properties and data publication on a large series of patients are of significant scientific interest and may be sensitive topics for the medical community, with possible divergent interest between medical societies, cooperative groups or other formal and informal national or European networks. Developers should have the option to choose between the different groups to optimise their good adhesion, the data quality and best costs management. Registry owners of patient registry should be independent from all products/ therapeutic utilisation, this independence would facilitate the governance and may help to find solutions for the sustainability issue. |  |
| **Sections 5.8.1.2 & 5.8.1.3****pg 30**  |  | Both these sections provide a useful description of the roles of industry and regulators, but neither paragraph appears to be about governance, beyond the opening paragraph of 5.8.1.3 |  |
| **Good registry practice**5.8 GovernancePg 30, line 8  |  | What should these research contract templates entail? |  |
| **Good registry practice**5.8 GovernancePg 30, line 11 |  | *Proposed change*: **Consider including the possibility to use registries not only for the post-authorisation monitoring of the products, but also in the pre-authorisation setting as a support to clinical development**.Unless a Qualification Opinion is obtained, in the absence of some broad “qualification process” for registries, there will always be a need to conduct individual feasibility– which can be resource-consuming and potentially duplicative across companies. **Once feasibility is complete – could this be made publicly available, to prevent the need to re-do?** |  |
| **Good registry practice**5.8 GovernancePg 30, lines 17-20 |  | ***Role of pharmaceutical companies***to initiate discussions with disease registry coordinators and regulators before –using the scientific advice procedure- or at an early stage of a marketing authorisation application on the relevance and adequacy of one or several existing patient disease registries for the long-term monitoring of their specific product;Scientific Advice Board may modify proposals for improvement but not prohibit ethically meaningful research questions for a pharmaceutical company. |  |
| **Good registry practice**5.8 GovernancePg 30, lines 25-6 |  | Joint protocols are complex and time consuming, any centralisation at regulatory level would help. This should not be mandatory, as there may be call for flexibility for given registry requirements. Getting this information is not easy. If the registries follow the recommendations made above, then this might be easier.Proposed change (if any): To understandthe extent and detail of data available…  |  |
| **Good registry practice**5.8 .1.1 role of registry coordinatorsPg 30-31 |  | Data sharing is almost always impossible if not prohibited per registry governance. Hence we fully support further guidance on data sharing and data ownership and consider that **the report should encourage more directly** the various stakeholders to accept the sharing of pseudonymized patient-level data, as outlined in a further comment below. |  |
| **Good registry practice**5.8 .1.3 role of regulatory authorities, lines 6-7\*\* |  | Within the scientific advice procedure, opinion for the registry study protocol can be sought. Having a prospective endorsement on study protocols and statistical analysis plans like for PASS studies by PRAC would greatly facilitate the use of registries in other regulatory settings.**Proposed to add the following bullet point before line 6:*** to review registry study protocols and statistical analysis plans and endorse those prospectively

Comment: Some of these only apply to new registries. For example, it will be hard for regulators to provide guidance to some of the existing registries.Proposed change (if any): to provide methodological guidance **for new registries** on core data elements… |  |
| **Good registry practice**5.8 GovernancePg 30, line 33 |  | *To support medicines evaluation …* Also to guide public health decisions and surveillance (if applicable). |  |
| **Good registry practice**5.8 Governance5.8.2 Data OwnershipPg 31 lines 11-29 |  |  The title “data ownership” is not correct; the issue of “ownership” is irrelevant and rather technical and complicated one to be addressed in this paper. What is important is to identify the privacy elements to ensure full compliance with GDPR. The term “data privacy” would be more appropriate. Sections 5.8.2. and 5.8.4. seem to presuppose that the ***only*** legal basis available for the processing of personal data within the context of a registry is the prior consent of patients. We are of the opinion that further consultation is required to reach a conclusion on that topic; it is advisable that the Cross-Committee Task Force on Patient Registries seeks to receive the advice of the European Data Protection Board which has recently published a detailed Opinion on a related concept, i.e. the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (Opinion 3/2019). In its Opinion 3/2019, the EDPB has successfully presented the limitations generated from a potential reliance to “consent” as a legal basis for data processing within the context of a clinical trial. Similar limitations may be applicable to the case of disease registries. As a result, it is imperative to initiate a discussion with a view of establishing a position whether another legal basis (and not consent) shall be the appropriate legal basis for data processing activities within the framework of disease registries. Given that the types and purposes of disease registries are not homogeneous (see e.g. section 3.2 of the Discussion Paper), it shall be always an obligation for the data controllers, i.e. the owners and coordinators of any disease registry, to identify their appropriate legal basis. For example, there could be cases of disease registries, where the processing of personal data by data controllers could be considered as *“necessary for the performance of a task carried out in the public interest”* pursuant to Article 6(1)(e) GDPR. Article 6(3) GDPR further provides that this basis shall be laid down by Union or Member State law and that the purpose of the processing shall be laid down in that legal basis. There may be other situations where the processing of personal data could be *“necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject”* following Article 6(1)(f) GDPR. Depending on the specific circumstances of a disease registry, the appropriate Article 9 condition for all processing operations of sensitive data could either be *“reasons of public interest in the area of public health [...] on the basis of Member State law”* (Article 9(2)(i)), or *“scientific ... purposes in accordance with Article 89(1) based on Union or Member State law”*(Article 9(2)(j)).The Discussion Paper makes no reference to the need to conduct a Privacy Impact Assessment (PIA) before initiating any disease registry. Within the framework of such PIA, the registry owners/coordinators shall determine, among others, the legal basis upon which the data processing activities shall be performed. The above reveal a need for further discussion and analysis that will require a separate guidance to be developed subsequently, and this perhaps is what needs to be flagged in this section.*Proposed changes:** Integrate sections 5.8.2. and 5.8.4. under a new section with the following heading: “Data Privacy”.
* The following could be added here: “It is the responsibility of local centres and registry coordinators to process and share personal data in accordance with the EU General Data Protection Regulation (GDPR). They shall conduct, where necessary, a Privacy Impact Assessment to consider all elements and risks of the underlying personal data processing activities including determining the appropriate legal basis for such activities. Further consultation is required to give guidance in relation to the options available for the legal bases that could be used and relied on for data processing activities in relation to disease registries. In situations where data are shared with MAHs based on a contractual agreement, the research contract should clearly describe the level of access to data, the intellectual property rights arising from the use of the data and the dissemination of the results. In case of specific studies performed by registry investigators and funded fully or partially by MAHs/MAAs, it is recommended to refer to the ENCePP Code of Conduct (15) (and ADVANCE Code of Conduct for vaccines (28)) in the research contract to ensure scientific independence and transparency, while allowing sharing of unpublished results with regulators and the MAHs/MAAs concerned. Local centers and registry coordinators shall ensure that in all cases, they provide to patients the appropriate privacy information. Patients need to be aware of why data is collected, what data is collected, how it will be used, and by whom and with whom it will be shared, and at what level of details. In addition, some patient registries have been expanded to include additional data such as genetic profiling and other biochemical analyses. This data is sensitive and it is important that patients have a good understanding of the data that could be provided to external organisations. Principles of informed consent, where consent is applicable, should be applied in accordance with the GDPR (30).”
 |  |
| **Good registry practice**5.8 Governance5.8.2 Data OwnershipPg 31 lines 22-29 |  | **Data ownership and intellectual property**Some registry coordinators have expressed concerns that unpublished data communicated to regulators or to MAHs in the context of regulatory procedures could be made publicly available through Regulation (EC) No 1049/2001 regarding public access to documents, thereby undermining their intellectual property rights. However, as highlighted in the EMA policy on access to documents (29), the EMA may restrict the access to documents and, prior to the release of documents, will always liaise with the entity having produced the document (in this case a registry coordinator or a MAH) to discuss potential steps necessary to protect commercially confidential information and personal data if applicable.Registries may and will contain unpublished efficacy and safety data – if registry does not belong to MAH it is inevitable to inform and share data with MAH and grant the possibility to take a stand before publication or considering the data for regulatory purposes; **Proposed change:** The use of unpublished data in that respect needs to be discussed in more detail. Create a separate paragraph or section with the heading “Intellectual Property”.  |  |
| **Good registry practice**5.8.3 Data sharingPg 31, line 35-37 |  | “Aggregated data (supported by statistical analysis if needed) are generally sufficient. Regulators will rarely request patient-level data or analytical datasets, and requests for pseudonymised patient data would be based on important public health reasons. Sharing of registry data should be based on a voluntary agreement between different parties.”The text is very ambiguous and will create more uncertainty than provide guidance. It would be really important if further guidance could be given when and mainly how data shall be shared between various stakeholders. This topic could be linked with the Anonymisation Project that EMA is currently running. |  |
| **Good registry practice**5.8 Governance5.8.3 Data SharingPg 32, line 16 - 19 |  | External control cohort can also be based on patients which are treated, but in the context of clear standard of care. So the statement on “unexposed patients” should be minimised.*Proposed change*: To complement with “data on unexposed patients to the experimental drug”.Only if the registries meet quality standards, which are not clearly documented. |  |
| **Good registry practice**5.8 Governance5.8.3 Data sharingPg 32, line 28 |  | This could imply that the option for amending the PSUR schedule is specific to circumstances where a registry is in use whereas in fact this could be required under other circumstances.*Proposed change*: **As with PSURs that do not include reference to registries, this** schedule may be amended depending on circumstances; |  |
| **Good registry practice**5.8 Governance5.8.3 Data sharingPg 33, line 1-12 |  | “The range of situations where MAHs may contact registries to get data is therefore potentially very large and each situation may require different agreements for data sharing, for example provision of statistical reports generated by the registry, of electronic tables of study results to be included in a study reports, of pseudonymised analytical datasets including only the data elements to be analysed, or of study reports to be submitted by companies to regulators. Provisions of pseudonymised patient-level data or analytical datasets to companies are generally not accepted by registries for policy reasons or concerns about patient privacy. In such cases, data analysis might be performed either by the registry team or by a trusted third-party.”Again, there is no reference to anonymization techniques that could be used and the guidance is lacking. The last sentence, which provides for a third party acting as the intermediary analysing data and thus having access to data instead of the MAH, requires further elaboration: what type of contractual provisions will be in place, what safety measures used, how this will affect the legal basis for the underlying processing, allocating the appropriate roles (controller vs. processor) between the parties, etc..  |  |
| **Good registry practice**5.8 Governance5.8.3 Data sharingPg 33, line 27-28 |  | If the regulators request an assessment of the effectiveness of risk minimisation measures it would be useful for the MAH to be notified if this were to impact an RMP put in place by the MAH. |  |
| **Good registry practice**5.8.3 Data sharingPg 33, lines 13 – 30 |  | This section addresses sharing with MAA/MAHs and Regulators. The Regulators section could be interpreted to also include health technology assessment bodies (HTABs), but this should be made explicit. However, it would be more appropriate to include a separate section for HTABs. This would be in line with the reference in Section 4, pages 16-17, lines 41-44, 1-2. *Proposed change:* Include a section 5.8.3.4. to address Data sharing with HTABs. |  |
| **Good registry practice**5.8.4 Informed ConsentPg 33, lines 31-33Pg 34 lines 12-15 |  | The requirement that “patients need to be aware of why data is collected, what data is collected, how it will be used, and by whom and with whom it will be shared, and at what level of details” in line 32-33 seems slightly contradictory to lines 12-15 on page 34 stating “If a new informed consent form is developed or if the current consent needs to be amended, it should be ensured that the **informed consent is broad enough to cover all potential uses of registry data in line with the applicable legislation, including the option for data sharing/pooling between registries and across country borders and with other stakeholders including regulators and MAHs”.**This contradiction highlights the limitations inevitably generated if consent is regarded as the only legal basis available. Under GDPR, consent shall be, among others, “specific”; it will always be a challenge to meet the requirements of granularity when consent language used is “broad enough” as the Discussion Paper suggests.In any case, please see above a proposed change and new wording. The report can also provide transparency on how data will be used in future by means other than the consent form.*Proposed change*: Suggest rewording as follows: “If a new informed consent form is developed or if the current consent needs to be amended, it should be ensured that the informed consent is broad enough to cover all potential uses of registry data in line with the applicable legislation, including ~~the option for~~ **explaining the need for any** data sharing/pooling between registries and across country borders and with other stakeholders including regulators, HTA bodies and MAHs”.Could provide a recommended (but not mandated) template or checklist of expected contents of the ICF. Clarification is required regarding when informed consent is required e.g. collecting data in a registry, conducting a registry study. Also how to collect patient ICF for retroactive use of the data?The section about informed consent implies that consent would be the legal basis for the processing of data. However, this may not be the only possibility, and for processing of data that are done for research purposes or to evaluate safety of pharmaceutical products and/or treatments it may be more appropriate to use other legal grounds such as those found in GDPR, Article 9 i and j.The downside of relying solely on consent is specifically for situations, where data must be retained for the verification of any claims that may have been made in scientific reports or publications, is that a withdrawal of consent from the individual means that the data should be deleted or fully anonymized according to the Article 29 working party guidance on consent (17/EN WP259 rev.01.), section 7.2 as well as the European Data Protection Board Opinion 3/2019 concerning the Questions and Answers on the interplay between the Clinical Trials Regulation (CTR) and the General Data Protection regulation (GDPR). So in case there is a need to be able to verify any research that are made based on data in a patient disease registries it is more appropriate to rely on other legal grounds (e.g. GDPR, Article 6(1)(e) or 6(1)(f) and Article 9(2)(i) and Article 9(2)(j)) than consent, and in case a consent mechanism is used it may be more appropriate to view that as an “additional safeguard” as suggested in the 29 working party guidance on consent (17/EN WP259 rev.01.) rather than a legal basis for the processing. |  |
| **Good registry practice**5.8.4 Informed ConsentPage 33, line 33-36 |  | Please consider changing the phrasing, since not only the genetic information is sensitive!Proposed change (if any): “The health data are sensitive and it is important.....” |  |
| **Good registry practice**5.8.4 Informed ConsentPg 34, lines 12-15 |  | **If a new informed consent form is developed or if the current consent needs to be amended, it should be ensured that the informed consent is broad enough to cover all potential uses of registry data in line with the applicable legislation, including the option for data sharing/pooling between registries and across country** **borders and with other stakeholders including regulators and MAHs.** Informed consent form: sharing data across country borders is difficult if data is shared outside EU due to different data protection laws etc.; oversight for data security and its usage may also be difficult. The registry owner needs to have appropriate governance in place to ensure that they always are in control of the sharing, and any potential further onwards transferObtaining consent can be very difficult in some circumstances (e.g. patient dying shortly after diagnosis) which can lead to selection bias. In countries in which waivers of informed consent for such circumstances are not possible, this could be acknowledged as a limitation. |  |
| **Good registry practice**5.8.4 Informed ConsentPg 34, lines 17-20 |  | **Data security**Security measures should be implemented to maintain the privacy of patients enrolled in a registry, described and documented in standard operating procedures. Security measures should be applied for data collection, storage, archiving, transmission and access. Each registry has the responsibility to ensure the security of the data collected in line to the provisions of the GDPR (30). **Proposed change:**Instead of “security” measures, reference shall be made in the heading and in the text to the appropriate “technical and organisational measures”. Please add information on who does the oversight.This shall be linked to the PIA which shall be conducted for registries. Additional guidance on that topic would be welcomed. Are there any minimum technical and organisation measures that could be applied across the board? |  |
| **5.8.5 Data security** |  | Suggest addition as follows ""to the provisions of the GDPR and other applicable EU and national legislation" |  |
| **Section 6: REGISTRY STUDIES** |
| **Registry studies**6.1 Registry contextPg 35, lines 22-24 |  | Clarification should be provided regarding how long analytical datasets and statistical programmes, used for generating the data included in the final study report, should be available for auditing and inspection (especially if kept by an external registry).*Proposed change*: add number of years this information should be kept available. |   |
| **Registry studies**6.1 Registry contextPg 35, lines 34-36 |  | Clarification is requested regarding whether study information can be entered by one MAH only in the case of joint study. |  |
| **Registry studies**6.1 Registry contextPg 35, line 37 |  | The abbreviation PASS was introduced on page 35, line 7. Please consider doing the same for the abbreviation PAES. |  |
| **Registry studies**6.2 TimelinesPg 36, lines 9-10 & 16 |  | There is mention of contract agreement per site, centre identification and individual centre. It is expected that the contract would be with the registry coordinator on behalf of participating centres. Rules in terms of data availability and quality are likely at the registry level.*Proposed change*: Delete any reference to individual centre and replace “contract agreement per country and site” by “per country and registry”.  |  |
| **Registry studies**6.3 Study protocolPg 36, line 25 |  | The protocol should not only ensure the internal validity of the study, but also the external validity.*Proposed change*: Refer to both internal and external validity of the study. |  |
| **Registries studies\_****6.3 Study protocol****Pg 36, line 29-32** |  | Definition alignment with GVP modules VI and VIII:This paper states “…secondary data collection, where the data for the study are already available and extracted from a dataset, and primary data collection, where the events of interest for the study are collected directly from patients as they come to the attention of the investigator.”GVP modules VI and VIII also include in the definition for secondary use of data “data previously collected from patients and healthcare professionals for another purpose.”The concept of the initial purpose of data collection (whether for the study or for another purpose) does not seem to have been implemented into the definitions of primary and secondary use of data. Moreover, the paper at times appears to conflate collection and extraction, and this could be better distinguished. |  |
| **Registry studies**6.3 Study protocolPg 36, lines 35-37 |  | How to handle when individual drugs are approved at different times and the protocols are therefore committed to on a rolling basis?Proposed change (if any): …to perform a study, **when possible,** MAHs are encouraged to design a joint registry study. |  |
| **Registry studies**6.3 Study protocolPg 37, line 2 |  | A sample size that is attainable may be set, but it may not be a sample size powered for anything meaningful.  It’s a two-step process: 1) figure out what sample size you need for meaningful studies, 2) determine if that sample size is attainable.  This is not clearly delineated here.  See previous comment on the executive summary on the use of the term conservative assumptions.*Proposed change*: Please consider replacing with “realistic”. If using “conservative”, please define the term (e.g. as a footnote or in glossary). |  |
| **Registry Studies****6.3 Study protocol**Pg 37, line 8 |  | For imposed PASS, the final protocol should be endorsed by PRAC before study start. For non-imposed studies included in a RMP, the protocol may also need to be agreed by regulators upon request from PRAC or CHMP. For clarity and consistency, consider aligning the terminology to GVP module V and VIII: Use “required in the risk management plan” (category 3), instead of “non-imposed”.  |  |
| **Registry studies**6.4 Study populationPg 37, lines 15-17 |  | The use of “may be” in the sentence is not appropriate as for studies based on registry, patients will be extracted from the registry/-ies. Additional data may be collected but a certain amount (or all) the data will come from the data already collected in the selected registry/-ies.*Proposed change:* Consider to replace “may” with “will”, and delete “depending of the objectives of the study”. |  |
| **Registry studies**6.4 Study populationPg 37, lines 20-23 |  | Use of the terms incident and prevalent may be confusing as these terms can be defined based on a disease or a treatment start. Especially as prevalent refers to the use of the drug and not the disease start in other sections such as lines 44-45 on page 39. *Proposed change*: Consider adapting to make the distinction clearer between incident diagnoses and incident drug initiation.  |  |
| **Registry studies**6.4 Study populationPg 37, lines 25-27 |  | This appears to be a recommendation - but prevalent user bias is discussed on pg 39.While this recommendation to include both prevalent and incident drug users in a registry is easily understandable, for many registry studies (in particular those involving treatment comparisons) prevalent users greatly increase the risk for bias (as alluded to in section 6.7, lines 43-11). This recommendation therefore reads strongly. **Proposed wording:**“The choice of including incident or prevalent patients in the study population has important implications for the data analysis and the interpretation of the results (see section 6.7). It is therefore recommended to collect the data needed to distinguish incident and prevalent patients and decide on whether to include incident and prevalent patients based on the study objectives and bias-variance trade-offs.” |  |
| **Registry studies**6.4 Study populationPg 37, lines 27-29 and 31-33 |  | This part refers to both patients included in clinical trials and patients coming from observational settings. Patients from clinical trials could also be considered as new users to some extent (similar comment as for the executive summary regarding the use of patients coming from clinical trials in a disease registry “later on”).*Proposed change*: “… could also be enrolled in a disease registry **post-clinical trial** ~~later on~~.”  |  |
| **Registry studies**6.4 Study populationPg 37, lines 30-31 |  | The use of the term “prospective recruitment” may be misleading, as by definition all patients within a registry are prospectively recruited. It should be clarified that this is related to the situation where additional data collection is needed specifically for the study purpose. *Proposed change*: The term “sequential inclusion” could be replaced by “appropriate inclusion”. |  |
| **Registry studies**6.4 Study populationPg 37, lines 33-38 |  | Would this need to be done in all cases? **A retrospective study alongside prospective studies is not always planned.** |  |
| **Registry studies**6.4 Study populationPg 37, lines 36-38 |  | There is a confusion between the registry set-up and the design of the study based on the registry data, where patients will be included following inclusion and exclusion criteria as defined in the study protocol. Some of these patients may need to be contacted so that further data are collected. The term “recruited” may not be appropriate to define these patients.*Proposed change:* Please consider clarifying that these patients are not recruited for the study itself, as they are selected in the registry/-ies, but that some of them can be contacted for additional data collection. Then, it makes sense to ensure that the patients who will accept the additional data collection will not differ from the ones who refused. |  |
| **Registry studies**6.5 Data collectionPg 37, lines 39-44 |  | Event or AE adjudication may require additional resources/expertise, which is not mentioned. |  |
| **Registry studies**6.5 Data collectionPg 37, lines 40 |  | There is a need for a similar discussion about the potential legal bases that could be relied on for the collection and processing of personal data within the context of registry studies (see comments in relation to section 5.8.2. and 5.8.4.). It may be under specific circumstances that consent is not the appropriate or the only legal basis for data processing activities. |  |
| **Registry studies**6.5 Data collectionPg 37, lines 42-43 |  | The word “collect” in this context may be misleading.*Proposed change:* Replace by “extract”. |  |
| **Registry studies**6.5 Data collectionPg 38, lines 29-30 |  | “The fact that a registry study is considered interventional has no influence on the non-interventional nature of the registry itself.” The sentence is confusing as the discussion paper intents to distinguish both. The statement appears only to be relevant for a registry study as part of a registry.*Proposed change*: **The fact that a registry study is considered interventional has no influence on the non-interventional nature of the ~~registry~~ data collection itself**.A reference to describe the difference between interventional and non-interventional registry studies would be useful. |  |
| **Registry studies**6.6 Data qualityPg 38 |  | This paragraph is not completely adapted to data quality control for registry studies but includes information about the registry data quality. For a registry study (not a registry), there is no source data verification against the source document. Source data verification is done on the extraction, to check the validity of the extraction process.*Proposed change:* Consider referring to the relevant section when dealing with registry data quality and not registry study data quality.The text indicates that source data verification would be expected in individual registry studies conducted for regulatory purposes. For that purpose, it would be a concern. if source data may be deleted due to the withdrawal of the consent from patients. i.e. it is not possible to verify the validity of the research in case consent is the sole legal ground form the processing.Instead of heavily relying on consent it is recommended that the legal basis for the processing of personal data be aligned with the European Data Protection Board Opinion 3/2019 and, if considered necessary further look into the governance and safeguards for a patient registry.Therefore, the context of the SDV should be clarified. Does it apply only to prospective patient identification for a specific registry study, or all registry data whether it is secondary use of data or not? What if the site is using a registry software similarly to an EMR; should the registry data be considered to be source data?Regarding the requirement to perform SDV, we recommend to refer to the EMA Reflection paper on risk based quality management in clinical trials (EMA/269011/2013) to **adopt similar risk-based methodologies and measures for registries based on the individual level of risk of each registry**. The focus should be on remote quality control measures and using targeted visits and targeted SDV triggered by pre-defined thresholds of data quality measures for high risk outcomes data only |  |
| **Registry studies**6.6 Data qualityPg 38, lines 33-43 |  | The need for additional measure for data quality control in a registry study depends on the measures applied routinely in the registry. In order to ensure acceptable data quality for individual registry studies conducted for regulatory purposes, source (registry) data verification and periodic auditing on a reasonable amount of data may need to be conducted on a risk analysis-based approach and following a strategy dependent on the scope of the study. As a general rule, source data verification for a minimum of 10% of randomly selected patients registered in individual study centres would be considered adequate. The level of data verification will have to be agreed upfront between the registry coordinators and the MAHs in the context of the study performed. Appropriate measures in case relevant findings are observed should be specified. Quality of short- and long-term data should be assured. Agreements on relevant logistical aspects should be made between the registry and MAHs in advance of study start.**Proposed change:**Please clarify how to make quality measures transparent and verifiable for 3rd party requests (eg MAH working with registry data from an independent registry). |  |
| **Registry studies**6.6 Data qualityPg 38, lines 37-39 |  | See previous comment regarding the **threshold of a minimum of 10%** for source data verification.10% may be relatively low or high, depending on the study population and frequency. For rare disease medicines, the 10% may be low, but for indications with a large number of patients, it is ambitious. This figure may have been taken from a routine monitoring sample selection and may suggest that it is standard.*Proposed change: S*uggest deleting the “threshold of a minimum of 10%”. **The next sentence gives flexibility to use a risk-based approach**. |  |
| **Registry studies**6.6 Data qualityPg 38, lines 39-43 |  | The use of MAH here is too restrictive as a registry study could be also with other stakeholders.*Proposed change*: Please consider to be more inclusive and not limiting to MAHs in this context. |  |
| **Registry studies**6.7Data analysisPg 39, lines 1-45 |  | How is data analysis handled if registry data is utilized to provide ‘historical’ controls to clinical trial data and analysed as one data set? Or in case of network meta-analysis?Can MAHs be involved in the third-party selection (what if adequate data collection at registry level but few methodological resources?) Likewise, can MAHs be involved in developing statistical analysis plans (line 7-9)? |  |
| **Registry studies**6.7 Data AnalysisPg 39, Lines 11-12 |  | Comment: The data are often already generated, in the case of existing registries.   Proposed change (if any): Pre-specification of the analytical approach before data are **tabulated in an analytic way** on any investigational or… |  |
| **Registry studies**6.7 Data AnalysisPg 39, Lines 18-25 |  | This section addresses “event rates” and “incidence rates.” It is recommended that the potential need to consider person-time of observation, as opposed to just number of people, also be included in this section.*Proposed change*: Comparisons between categories should take observation periods into account and use parameters such hazard ratios (HR) or incidence rate ratios (IRR) adjusted for potential confounders. **The potential need to consider person-time of observation, as opposed to just number of people, should also be considered.** |  |
| **Registry studies**6.7 Data AnalysisPg 39, lines 26-27 |  | The term ‘systematically’ may be too strong in this context, as there may be some situations (even if not the rule) where the differences in patient groups given different treatments to be compared in a study may be minimum.*Proposed change:* Delete the term “systematically”. |  |
| **Registry studies**6.7 Data AnalysisPg 39, lines 26-42 |  | **Propensity scores:** Defining the use of the propensity score approach is somehow restrictive, it is suggested to reword this in more general language.*Proposed change*: This potential bias can be addressed using various methods e.g. the propensity score approach. The appropriate method and the assessment of whether the propensity score approach is valid will depend on the research question, patient population, outcomes for the specific study and data available. Of note, there are additional matching methods available in the literature, these methods (e.g. Entropy Balancing (Hainmueller 2012) or Genetic matching (Diamond 2013) potentially outperform propensity scores in bias reduction. |  |
| **Registry studies**6.7 Data AnalysisPg 40, lines 10-11 |  | The last sentence of the paragraph is not clear. Indeed, in a comparative safety or effectiveness study, the objective will be to compare new users of the newly marketed treatment, with new users of one or several competing treatments, to avoid the biases as listed in the section.*Proposed change*: Consider deleting or rephrase to clarify. |  |
| **Registry studies**6.7 Data AnalysisPg 40, lines 23-29 |  | If using a comparative non-exposure group, particularly if using data from a different country or region, care should be taken to ensure that underlying differences in the two samples/populations which may influence the likelihood of outcome occurrence can be adequately measured and controlled for in the analysis. If not, additional confounding may be introduced. For example, populations in different regions may have significantly different underlying risks for an outcome due to genetic or environmental differences which may distort results if not identified and controlled for.The document should acknowledge that after product approval, it is important to understand the possibilities and limitations of comparators and the associated bias/confounding in a non-randomized setting: appropriate planning with epidemiology expertise is crucial Should the list of considerations include the risk of systematic differences between the exposure group within the registry and the group outside the registry? Especially if countries within and outside are very different in terms of life expectancy or disease management.*Proposed change*: Suggest adding text to recommend accounting for underlying population differences to the extent possible if using a comparative non-exposure group. |  |
| **Registry studies****6.8 Safety Reporting**Pg 40 |  | In alignment with GVP modules VI and VIII, consider adding the following guidance to this section: *For combined study designs, the same requirements as for studies with primary data collection should be followed for adverse events obtained through primary data collection and the guidance for studies with design based on secondary use of data should be applied to adverse events based on secondary use of data*. |  |
| **Registry studies****6.8 Safety Reporting**Pg 40 lines 34-38 |  | Methods to reduce the problem of missing data could be very complex and have strong assumptions (e.g. Multiple Imputation by chained equations). **Proposed change:** Include a link/reference to a common guideline which describes the application of methods to handle missing would be very helpful. |  |
| **Registry studies**6.8 Safety reportingPg 40, lines 40-42 |  | Same comment as above on the use of design based on primary data collection, which should only be used to complement registry data in this context of registry-based studies.*Proposed change*: Consider to adapt to make it more clear. |   |
| **Registry studies**6.8 Safety reportingPg 40-41 |  | One of the operational challenges faced by MAHs when there is a need to collect additional data (e.g., specific safety data) is that this additional data collection is often **done via an additional module in the registry** (as stated in section 6.5). Then, Time 0 (start date) to be considered for the start of the reporting procedure is a challenge, as often cannot be when the centre enters the data into the additional module of the registry. Should it be when the registry coordinator is aware of the information? This should be clarified. In addition, the agreement is between the MAA/MAH and the registry (coordinator) and not with all centres participating to the registry, even in case of additional data collection needed.*Proposed change:* To further describe the process to apply within the specific setting of the additional data collection within the registry. |  |
| **Registry studies**6.8 Safety reportingPg 41, line 5 – 7 and 28 - 31 |  | No reference to MedDRA.Is the following guidance from GVP modules VI and VIII applicable to this registry guidance regarding primary collection of data?“*Any reference to adverse events that will not be collected should be made using the appropriate level of the MedDRA classification.”* |  |
| **Registry studies**6.8 Safety reportingPg 41, line 21 |  | This line, taken from GVP Module VI, refers to adverse events specified in the study protocol which are **not** systematically collected and may be reported as spontaneous adverse reactions where appropriate. More often the protocol specifies which adverse events **are** to be systematically collected.*Proposed change*: Where a study protocol specifies only those adverse events which are to be systematically collected, please clarify how safety reporting should be managed for any other adverse events/reactions incidentally identified with the product within the same study. It is unclear whether these should also be reported as solicited events, with causality assessed accordingly or if they should be handled as spontaneous reports, where appropriate. |   |
| **Registry studies**6.9 Reporting of study resultsp. 42, lines 1-3 |  | Would this be considered the responsibility of the MAH or registry coordinator?**Proposed change:** We propose registry coordinators as he/she is the one most likely training the site and communicating with the centres/investigatorsWho is an investigator? Is this the registry coordinator (page 43, line 15)? Individual HCPs? It is not clear how this can be done (e.g., if registry study is initiated by MAH using the data already in the registry)**Proposed change:** We propose to use consistent wording already in the concepts defined in section 3. In that case registry participants? |  |
| **Registry studies**6.9 Reporting of study resultsp. 42, lines 20-21 |  | Who would be the “lead investigator” in this context?Wouldn’t this be exactly what registry coordinators are concerned about, as described on page 31, line 22-25. Would this not hence contradict the proposed approach by the EMA regarding intellectual property (page 31, line 25-29)?Focus is on the rights of the coordinator (or lead investigator) for publication but nothing is mentioned regarding the rights of the MAH.**Proposed change:** Define lead investigator or change to principal investigator (if applicable). |  |
| **Registry studies**6.9 Reporting of study results |  | Inclusion of MAH in publication should be based, if applicable, on relevant authorship guidelines (references to ICMJE to be added)It is stated that the MAH should have the ability to comment on draft publications. Presumably, final decisions on content are made by the investigator, not the MAH. That should be made clear.*Proposed change*: **The MAH which funded the registry study should be entitled to view the final results and interpretations thereof prior to submission for publication** and to comment in advance of submission within a reasonable time limit, e.g. one month, and without unjustifiably delaying the publication. **While the final decision on content should be made by the lead investigator, t**he MAH may also require that the presentation of the results be changed to delete confidential information. |  |
| **Section 7: CONCLUSIONS** |
| **Conclusions**Pg 43, lines 7-9 |  | This is important information not only for the pharmaceutical companies but also regulators. Also registries are already being used by industry.*Proposed change:* Include “…to understand under which circumstances registries could be used by pharmaceutical companies and regulatory authorities to support…”. |  |
| **Conclusions**Pg 43, lines 20-26 |  | Challenges include ownership and accountability. Who will collect and manage disease registries and who will benefit are different entities, which may impact quality management and flexibility of implementing changes… Sustainability is also a challenge. |  |
| **Conclusions**Pg 43, lines 41-42 |  | “*The EMA is willing to support interactions and provide tools to facilitate recognition of disease registries as data sources to conduct studies for regulatory purposes*”However, there’s no mention of the minimum requirements of such registries in order to be acceptable, i.e. to provide scientifically robust data for regulatory purposes (e.g. PASS, PAES). |  |
| **Conclusions**Pg 43  |  | Despite opening with a great ambition – “good registry practice” the Conclusion takes a very narrow view. We are missing “next steps”, recognising that this methodological progress is still in motion and moreover that these requirements are international. It would be very welcome to see how this agenda could **be brought to international alignment through ICH.** |  |
| Page 44line 36-37 |  | Sustainability of disease registries has been a common issue discussed in the course of the EMA Patient registry initiative. Studies conducted with registry data may provide an additional source of funding from the public or the private sector. Governance principles are therefore proposed to facilitate interactions between all parties concerned while preserving the registry participants’ scientific independence. For this aspect, quality management is an important activity to provide confidence in the quality of the data that can be generated.The private sector should be involved as an equal player and possible contractual partner.**Proposed Change:**Studies conducted with registry data or registries conducted in public private partnerships can provide an additional source of funding from the public or the private sector. |  |

Please add more rows if needed.